

BRAND AND USABILITY IN CONTENT-INTENSIVE WEBSITES

Tao Yang

Submitted to the faculty of the University Graduate School
in partial fulfillment of the requirements
for the degree
Doctor of Philosophy
in the School of Informatics and Computing,
Indiana University

October 2013

Accepted by the Graduate Faculty, of Indiana University, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Mark S. Pfaff, Ph.D., Chair

Davide Bolchini, Ph.D.

Doctoral Committee

Karl F. MacDorman, Ph.D.

August 29, 2013

Anthony D. Cox, Ph.D.

© 2013

Tao Yang

Acknowledgements

It has been a long journey to arrive at this point. There are many people I want to thank who helped me start this journey, guided me along the way, and supported me when I encountered difficulties.

A special thank goes out to the former dean of our school, Dr. Darrell Bailey, and my advisor, Dr. Davide Bolchini. I first met Dr. Bailey six years ago in Guangzhou, China when I was working on my master's degree. He helped me apply for and become admitted to this great HCI PhD program and find the funding necessary to fully cover my studying and living expenses. Without his support, this whole story might not have happened. My advisor, Dr. Bolchini, was the person who had the most influence on my academic career. It has been a fantastic experience working with him and I always feel excited when thinking about how much I have grown during these years under his patient guidance and inspiring instructions. Every piece of this dissertation was imbued with his contributions and ideas and reading it brings back good memories of the wonderful and heated discussions and brainstorming sessions between me and him. I also want to thank Dr. Bolchini for his support in helping me correct my mistakes, especially in regard to the time when I opted to take an overly-advanced class at another department. Without his help, I might not have been able to overcome the difficulty so smoothly.

Many thanks to my research committee members: Dr. Mark Pfaff, Dr. Karl MacDorman, and Dr. Anthony Cox. Thanks go out to Dr. Pfaff for helping me enhance the experiment design found in the main chapter of my dissertation. Our meetings were short, but highly effective and I enjoyed reading the long emails that he sent me regarding the statistical tests that I needed to perform. Dr. MacDorman taught me most of the skills that I needed in order to conduct the empirical research within my dissertation. His classes were always challenging, but fruitful and the one-year rotation in his research lab vastly broadened my knowledge scope and enriched my skill sets. Moreover, I want to thank Dr. MacDorman for his help in regard to when I first came to the US and had problems fully comprehending the tons of reading materials in the classes. Thanks also need to go out to Dr. Cox for replying to my emails and agreeing to serve on my research committee when he did not know me at all. He had provided me with many ideas and

advice on the core part of my dissertation, which ended up with an acceptance to a prestigious HCI journal. I am also grateful for his help in regard to enrolling me in several, wonderful MBA classes.

My life in the US would not have been as wonderful as it has been without the company of my colleagues and friends. Thanks go out to Dr. MacDorman's students, Chin-Chang Ho and Himalaya Patel, for teaching me all the experimental design and statistical analysis skills. I always admire how knowledgeable they are in those fields. Thanks to my colleagues Romisa Rohani Ghahari, who is still pursuing her degree, and Mexhid Ferati, who is now a professor in Macedonia. The working and leisure times that we shared and all of the interesting conversations that we had will always be memorable to me. I would also like to thank Debaleena Chattopadhyay, Prathik Gadde, Yuan Jia, Arthur Liu, Sung Pil Moon, Wei Pan, and Afarin Pirzadeh for their help and ideas, all of which helped strengthen my dissertation. I wish you all good luck in your future endeavors and hope that you enjoy the rest of your time in the USER Lab.

Thanks go out to my master's advisor, Dr. Yongli Zou, for his continued support. I would like to thank him for introducing me to Mr. Gerry Sanders and his wife, Phyllis Sanders, who have taught me a lot about life and how to handle difficult situations. Thank you to all of my friends in China who kept on blessing me and supporting me, even though I am thousands of miles away from them.

Finally, I would like to thank my parents for bringing me up with love and respect, for tolerating all of the mistakes that I have made, and for allowing me to not be by their sides for so many years. They may never get a chance to read this dissertation, but I know they will be happy about it and proud of what I have accomplished.

Abstract

Tao Yang

BRAND AND USABILITY IN CONTENT-INTENSIVE WEBSITES

Our connections to the digital world are invoked by brands, but the intersection of branding and interaction design is still an under-investigated area. Particularly, current websites are designed not only to support essential user tasks, but also to communicate an institution's intended brand values and traits. What we do not yet know, however, is *which design factors affect which aspect of a brand*. To demystify this issue, three sub-projects were conducted.

The first project developed a systematic approach for evaluating the branding effectiveness of content-intensive websites (BREW). BREW gauges users' brand perceptions on four well-known branding constructs: *brand as product*, *brand as organization*, *user image*, and *brand as person*. It also provides rich guidelines for eBranding researchers in regard to planning and executing a user study and making improvement recommendations based on the study results.

The second project offered a standardized perceived usability questionnaire entitled DEEP (design-oriented evaluation of perceived web usability). DEEP captures the perceived website usability on five design-oriented dimensions: *content*, *information architecture*, *navigation*, *layout consistency*, and *visual guidance*. While existing questionnaires assess more holistic concepts, such as ease-of-use and learnability, DEEP can more transparently reveal where the problem actually lies. Moreover, DEEP suggests that the two most critical and reliable usability dimensions are interface consistency and visual guidance.

Capitalizing on the BREW approach and the findings from DEEP, a controlled experiment (N=261) was conducted by manipulating interface consistency and visual guidance of an anonymized university website to see how these variables may affect the university's image. Unexpectedly, consistency did not significantly predict brand image, while the effect of visual guidance on brand perception showed a remarkable gender

difference. When visual guidance was significantly worsened, females became much less satisfied with the university in terms of brand as product (e.g., teaching and research quality) and user image (e.g., students' characteristics). In contrast, males' perceptions of the university's brand image stayed the same in most circumstances. The reason for this gender difference was revealed through a further path analysis and a follow-up interview, which inspired new research directions to unpack even more the nexus between branding and interaction design.

Mark S. Pfaff, Ph.D., Chair

Table of Contents

List of Tables	xii
List of Figures.....	xiv
Chapter 1. Introduction	1
Chapter 2. Review of Theoretical Background.....	5
2.1. Modeling Brand in Communication and Marketing Research	5
2.2. Online Branding Strategies and Frameworks	10
2.3. Reciprocal Effects of Website Visiting and Brand Perception	11
2.4. Usability for Branding	12
2.5. Methods for Capturing Perceived Usability and Their Limitations	13
2.5.1. Universal Perceived Usability Questionnaires.....	14
2.5.2. Perceived Usability Questionnaires for Websites.....	16
2.5.3. Perceived Usability Questionnaires for Mobile Applications	17
2.5.4. Evaluation-Design Alignment Gap	17
Chapter 3. BREW: A Systematic Approach for Evaluating the Branding Effectiveness of Content-intensive Websites	21
3.1. The BREW eBranding Evaluation Approach	22
3.1.1. Phase One: Brand Attribute Inspection.....	23
3.1.2. Phase Two: Brand Perception Questionnaire Composition.....	26
3.1.3. Phase Three: Task Design	28
3.1.4. Phase Four: User Test	30
3.1.5. Phase Five: Result Analysis	36
3.2. Results of the Case Study.....	40
3.2.1. The Impact of the AT&T Website Itself	40
3.2.2. The Comparative Branding Effectiveness of the AT&T Website	42
3.3. From Evaluation Results to Strategies for Improvement.....	45
3.3.1. Strategies for Enhancing the Communication of the Attribute “Valuable”	46
3.3.2. Possible Improvement Strategies for Female Users’ Brand Perceptions	47
3.3.3. Strategies for Enhancing the Communication of the Attribute “Saving”	49
3.4. Discussions.....	52
3.4.1. Synopsis of the BREW Evaluation Approach.....	52

3.4.2. Online or Lab-Based Evaluation	53
3.4.3. Validity and Limitations of the BREW Approach	54
3.4.4. Validity and Limitations of the Case Study	55
Chapter 4. DEEP: Design-oriented Evaluation of Perceived Website Usability	57
4.1. Website Design Dimensions	57
4.1.1. Content.....	58
4.1.2. Structure and Information Architecture	58
4.1.3. Navigation	59
4.1.4. Labeling.....	59
4.1.5. Graphics.....	60
4.1.6. Technological Performance	61
4.2. Research Methods and Techniques	61
4.2.1. Creating Candidate Measurement Items Based on Design Dimensions	62
4.2.2. Performing Internal Ratings on the Candidate Items.....	63
4.2.3. Online Field Test 1: Revise and Validate Perceived Usability Dimensions	64
4.2.4. Online Field Test 2: Strengthening the Revised Perceived Usability Dimensions	67
4.3. Results.....	68
4.3.1. Results from the Internal Rating	68
4.3.2. Results from Online Field Test 1.....	69
4.3.3. Results from Online Field Test 2.....	82
4.4. Discussion.....	85
4.4.1. Perception of Labeling Versus Navigation	85
4.4.2. Perceived Technical Performance	85
4.4.3. Elements of Graphics Usability	86
4.4.4. Minimizing the Number of Items	87
4.4.5. Guidance for Using and Interpreting DEEP: The Designer's Perspective	87
4.4.6. Overall Value of the Study	90
4.4.7. Limitations	92
Chapter 5. Branded Interactions: Predicting Perceived Product Traits and User Image from Interface Consistency and Visual Guidance.....	94
5.1. Selecting the Most Reliable Design-based Usability Dimensions	94

5.2. Research Hypotheses	95
5.3. Methods	96
5.3.1. Choosing the Application Domain.....	96
5.3.2. Gauging Brand Perception	96
5.3.3. Experiment Design and Stimuli.....	98
5.3.4. Procedure of the Online Experiment.....	101
5.3.5. The Follow-up Semi-structured Interview.....	102
5.4. Results.....	103
5.4.1. Manipulation Checks	104
5.4.2. Brand as Product.....	106
5.4.3. User Image.....	109
5.4.4. Brand as Person.....	111
5.4.5. Investigating the Gender Difference on Brand Perception	111
5.4.6. Content Only and Full Website	116
5.4.7. Findings from the Follow-up Interviews.....	117
5.5. Discussion.....	124
5.5.1. No Effect of Consistency on Brand Image	124
5.5.2. Visual Guidance Affects Brand Image with Remarkable Gender Difference.....	125
5.5.3. Toward a Gender-sensitive, Dual-mediation Model of Brand Perception	126
5.5.4. Content Only Version Beat Full Website.....	128
5.5.5. Understanding Interface Design for Branding: Synopsis of the Contribution.....	128
5.5.6. Validity and Limits of the Study.....	131
Chapter 6. Summary of the Contributions.....	133
Chapter 7. Future Research Directions.....	136
Appendices	140
Appendix A. Brand Perception Questionnaires Used in BREW Case Studies	140
Appendix B. User Tasks for the BREW Case Studies	146
Appendix C. Demographics Questions.....	147
Appendix D. No Significant Difference Was Found Between AT&T and Verizon in the Perceptions of All Brand Identity Attributes after In-depth Website Exposure.....	148

Appendix E. Candidate Measurement Items Selected from Existing Questionnaires (49 Items)	149
Appendix F. Initial Questionnaire Created for Internal Rating	152
Appendix G. Perceived Usability Questionnaire Used in Online Field Test 1.....	156
Appendix H. Correlations of the Nine Usability Dimensions Extracted from EFA.....	160
Appendix I. Correlations among the Usability Dimensions and Success Rate after CFA	161
Appendix J. Perceived Usability Questionnaire Used in Online Field Test 2.....	162
Appendix K. DEEP: Design-oriented Evaluation of Perceived Usability.....	165
Appendix L. Correlations among Perceived Usability Dimensions, Overall Perceived Usability and Task Success Rate	167
Appendix M. Correlations among Perceived Usability and Aesthetics Dimensions	168
Appendix N. Website Experimental Stimuli	169
Appendix O. Fit Indices of the Path Models Illustrating the Visual Guidance * Gender Interaction	182
Appendix P. Follow-up Interview Findings – Why Users Gave a Particular Rating to a Specific Brand Identity Attribute	183
References	191
Curriculum Vitae	

List of Tables

Table 1. Eliciting AT&T's key brand identity attributes.	24
Table 2. Tasks used in online field test 1.	66
Table 3. Internal rating of the perceived usability measurement items.	69
Table 4. Factor and reliability analyses of the perceived content dimension.	70
Table 5. Factor and reliability analyses of the perceived structure and information architecture dimension.	71
Table 6. Factor and reliability analyses of the perceived navigation dimension.....	72
Table 7. Factor and reliability analyses of the perceived labeling dimension.....	73
Table 8. Factor and reliability analyses of the perceived cognitive effort.....	74
Table 9. Factor and reliability analyses of the perceived graphics dimension (original).	75
Table 10. Factor and reliability analyses of the perceived graphics dimension (revised).	75
Table 11. Reliability analysis of the perceived technological performance dimension.....	76
Table 12. Structural coefficients for the perceived usability dimensions.....	77
Table 13. Structural coefficients for the perceived usability dimensions (after revision).....	79
Table 14. Structural coefficients for the perceived usability dimensions (final).	82
Table 15. Guiding the interpretation of DEEP results – examples.....	87
Table 16. Exploratory factor analyses of the measurement items for consistency (C) and visual guidance (V) (N=362) (Yang, Linder, & Bolchini, 2012).....	95
Table 17. Operationalize university brand identity into more concrete attributes.....	97
Table 18. User tasks used in pre-tests and the actual experiment.	101
Table 19. Code consistency and visual guidance in two binary variables.	106
Table 20. Analyzing the visual guidance * gender interaction (* $p < .05$, ** $p < .001$; CRD: Critical Ratio for Differences; '=': no significant difference; '>': significantly higher for females).....	115
Table 21. Users' overall impressions toward the university (numeric data stand for the number of participants who gave a particular rating).....	118
Table 22. Number and types of reasons for particular ratings on specific brand identity attributes.	121

Table 23. Correlations between perceived consistency and each brand identity attribute (*p < .05; **p < .01).....	124
Table 24. A synopsis of the research findings (% Decrease: The percentage of the decrease in the brand perception rating when visual guidance was worsened).	129

List of Figures

Figure 1. Deconstructing usability and brand identity enables the study of design factors' effects on brand perceptions.	1
Figure 2. Aaker's (1996) Brand Identity Planning Model.	8
Figure 3. Addressing the Evaluation-Design Alignment Gap can help relate usability problems to the analytical composition of the system design.	18
Figure 4. An overview of the BREW evaluation approach.	23
Figure 5. Inspecting where the branding identity attributes are communicated within the website.	26
Figure 6. Realistic tasks allow researchers to examine whether users can naturally perceive the intended brand identity attributes.	29
Figure 7. The impact of a website on brand perception can be capture by comparing the brand perception of a testing group (i.e., in-depth exposure to the website) with that of a control group (i.e., no in-depth exposure to the website).	31
Figure 8. Using the OFTEN testing environment to conduct the BREW evaluation online.	34
Figure 9. When prior brand knowledge is measured by a single testing group, a mixed-model ANOVA can be conducted to examine the impact of website on brand perception.	36
Figure 10. When prior brand knowledge is measured by a control group, a between-subject ANOVA can be conducted to examine the impact of the website on brand perception.	37
Figure 11. Calculating the change in brand perception.	39
Figure 12. In-depth exposure to the AT&T website brought two significant changes in brand perception (*significantly different from the control group, $p < .05$).	41
Figure 13. The AT&T website may have a different impact on females and males in regard to brand perception (*significantly different from the control group, $p < .025$).	42
Figure 14. In-depth exposure to the Verizon website brought three significant changes in brand perception (*significantly different from the control group, $p < .05$).	43

Figure 15. The significant differences in brand perceptions between AT&T and Verizon diminished after in-depth website exposure (*significantly different from Verizon, $p < .05$).	44
Figure 16. The impact of the AT&T website on the perceptions of two brand identity attributes was different from that of the Verizon website (*significantly different from Verizon, $p < .05$).	44
Figure 17. With very small font and a not so important location, the attribute “valuable” is not emphasized enough on this page.	47
Figure 18. Visual guidance problems of the AT&T website may negatively affect females’ perceptions on several brand identity attributes.	48
Figure 19. A better visual guidance design by Verizon.	49
Figure 20. A synopsis of the BREW evaluation approach.	51
Figure 21. Hypermedia design dimensions play a pivotal role in structuring the requirements, design and evaluation activity (Bolchini & Chatterji, 2009; Bolchini & Paolini, 2004).	58
Figure 22. Procedure used to develop and validate the DEEP questionnaire: Identify candidate items → internal rating ($N = 12$) → online field test 1 ($N = 196$) → online field test 2 ($N = 362$) → final questionnaire.	60
Figure 23. By mapping the perceived usability to key web design dimensions, DEEP addresses the Evaluation-Design Alignment Gap.	91
Figure 24. The experiment websites were embedded into the OFTEN environment and tested through Amazon Mechanical Turk.	98
Figure 25. Examples of the experiment stimuli: ① a fictitious university name was used to control for prior brand knowledge; ② in the C–V+ version, the navigation bar was modified from horizontal to vertical to create inconsistent navigation; ③ in the C–V+ version, the color of the banner was changed to create inconsistent identity; ④ in the C+V– version, the pages were cluttered with poor background contrast and small fonts; ⑤ important content was placed where advertisements usually appear.	99
Figure 26. Successful manipulation checks of the website interventions. “*” indicates values that were significantly different from the C+V+ version ($p < .01$).	104
Figure 27. Worsened visual guidance (V) caused significant drop in B2 (Academic excellence) and B3 (Teaching facilities); but worsened interface consistency (C) did not seem to affect brand perception.	106

Figure 28. Females gave significantly lower ratings to all three teaching and research quality attributes when visual guidance was worsened, but the males did not (* $p < .05$).	107
Figure 29. Worsened visual guidance caused significant drop in B4 (Unique identity) and B6 (High ranking, <i>female only</i>); but worsened consistency did not seem to affect brand perception.	108
Figure 30. Females gave significantly lower ratings to B7 (Help students to plan for their future) and B9 (Offer a personalized path of study) when visual guidance was worsened, but the males did not (* $p < .05$).	109
Figure 31. Females gave significantly lower ratings to all three attributes about the personality of typical students when visual guidance was worsened, but the males did not (* $p < .05$).	110
Figure 32. In the presence of poor visual guidance, female users gave significantly lower ratings to all three attributes about the social skills of typical students (* $p < .05$).	111
Figure 33. A conceptual path model of the experiment.	112
Figure 34. Two-group path analysis on B1 (High-Quality Teaching); significant difference was found between β_{a3} and β_{b3} . (* $p < .05$, ** $p < .001$).	113
Figure 35. The pure content document yielded the most satisfactory perceptions on all brand identity factors (*significantly lower than the pure content group, $p < .05$).	116
Figure 36. A gender-specific dual-mediation model for website visual guidance and brand perception.	127

Chapter 1. Introduction

Our minds immediately evoke a world of experiences, expectations, values, and emotions simply by recalling or recognizing a popular brand. From Gap™, Apple™, Nike™, Google™, NFL™, Gucci™, to MIT, a dynamic ecosystem of brands pervades our lives to provide memorable and culturally shared landmarks in the complexity of the marketed world, both physical and virtual. Due to its pervasive and viral nature, web branding is used today as the fundamental vehicle by which to establish and reinforce a brand for most companies and institutions (Bolchini, Garzotto, & Paolini, 2008; De Chernatony & McDonald, 2003). Thus, shaping brand experiences clearly intersects interaction design on several crucial dimensions. By reinforcing a common assumption, leading user experience practitioners have argued that the reputation of a brand can be greatly contaminated by poor website usability (Baty, 2006; Spool, 1996; Spool, 2002). Marketing professionals have proposed guidelines and frameworks to lead effective online branding initiatives (Page & Lepkowska-White, 2002; Ries & Ries, 2000; Simmons, 2007). Both usability and brand, however, are highly complex constructs, which require a deeper analysis in order to determine how they influence one another. What we do not yet know is which design aspect of a website affects the communication of which specific brand attribute (Figure 1). With this knowledge in hand, HCI designers can greatly expand the impact of their design to achieve better brand communication.

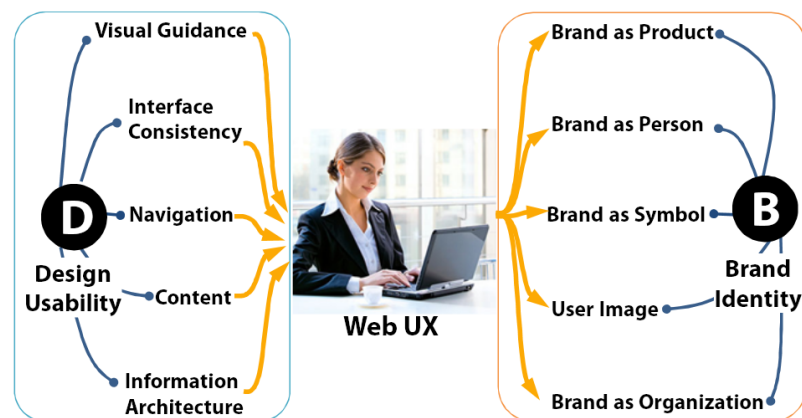


Figure 1. Deconstructing usability and brand identity enables the study of design factors' effects on brand perceptions.

In order to understand the influence of website usability on brand perception, however, I need to address two fundamental issues related to these two concepts.

First, existing eBranding strategies and frameworks do not cover a systematic evaluation method (Page & Lepkowska-White, 2002; Ries & Ries, 2000; Simmons, 2007). In other words, for an existing company or institutional website, how do we get to know its effectiveness in communicating the company or institution's key branding messages or attributes? For example, can Apple's website successfully communicate to the users its "think different" culture? Can Subway's website make users feel as if they want to "eat fresh" (i.e., Subway's slogan) at the restaurant? Without a systematic way to capture how well users perceive these attributes, I am not ready to explore whether and how users' brand perceptions are affected by certain website usability factors.

Second, existing perceived usability questionnaires do not capture users' perceptions on the usability of the analytical design factors of a website, such as the quality of the website's content, structure, navigation, and graphics (Yang, Linder, & Bolchini, 2012). If I want to manipulate the quality of some of these design factors to see their impacts on users' perceptions on certain brand attributes, I need to use a standardized questionnaire that can capture users' perceptions on these design factors in order to ensure that the manipulation is accurate (i.e., only the target factors are manipulated, other factors are not affected) and sufficient (i.e., the manipulation is noticeable by the users). The existing perceived usability questionnaires, however, can only be used to measure more holistic constructs, such as ease-of-use and learnability (Chin et al., 1988; Lewis, 1995; Kirakowski, Claridge, & Whitehand, 1998; Wang and Senecal, 2008), which do not provide information on the quality of the specific design-oriented factors.

Accordingly, this dissertation presents three research projects. The first project presents a systematic approach, known as BREW, used to evaluate the branding effectiveness of content-intensive websites. The BREW evaluation approach is built upon the Aaker's brand identity planning model and HCI user testing techniques. eBranding researchers can follow the guidelines suggested by BREW to plan and execute a comprehensive evaluation on the ability of an existing website to communicate key brand identity attributes and devise improvement or re-design recommendations and strategies based

on the evaluation results. The components and procedures of the BREW approach are illustrated by a case study on evaluating the branding effectiveness of the AT&T website. The second project depicts the full process of developing a standardized questionnaire to be used to measure the perceived website usability on five design-oriented factors: *content*, *structure and information architecture*, *navigation*, *layout consistency*, and *visual guidance*. Using this questionnaire, HCI researchers can directly learn which specific design-oriented aspects of a website need to be enhanced, as opposed to only knowing whether a website is easy to use as a whole. In addition, in the process of creating the questionnaire, I identified that, in addition to content, *interface consistency* and *visual guidance* are the two most reliable and uncorrelated website usability factors, which are good candidates for demystifying the relation between usability and branding. Based on this finding and part of the BREW evaluation approach, the third project presents a controlled experiment that investigates how the quality of the interface consistency and visual guidance of a US state university's website could affect users' perceptions of the image of the university. The results of the experiment reveal that interface consistency does not significantly predict brand image, while the effect of visual guidance shows a remarkable gender difference. When website visual guidance is significantly worsened, females become much less satisfied with the university in terms of brand as product (teaching, research quality, and student support) and user image (students' characteristics and social skills). In contrast, males' opinions toward the university stay the same in most circumstances. A follow-up semi-structured interview was carried out to further interpret the experimental findings.

The rest of the chapters in this dissertation are organized as follows. Chapter 2 reviews the theoretical background of this dissertation, which includes a discussion on the characteristics and shortcomings of the existing branding and eBranding models and frameworks, a summary of the related works on investigating the relation between website design and branding, and a review of the advantages and disadvantages of the current questionnaire instruments for measuring perceived website usability. Chapter 3 presents the components and procedures of the BREW evaluation approach and a case study on evaluating the communication effectiveness of the AT&T website. Chapter 4 documents the full process of developing a design-oriented perceived usability questionnaire, which consists of an expert rating session and two rounds of online field tests. Chapter 5 presents the controlled experiment and follow-up semi-structured

interview on investigating how interface consistency and visual guidance affect brand image. Chapter 6 summarizes the contributions of this dissertation, while Chapter 7 discusses possible future research directions.

Chapter 2. Review of Theoretical Background

This research required an in-depth understanding of both branding and website usability. Accordingly, in this chapter, I first review the existing models that illustrate the components or sub-constructs of a branding system, and the strategies that marketers suggest to take in order to achieve the desired branding effects through the online channel. Then, in order to understand in what ways and how much the relationship between website usability and brand perception has been investigated, I review and analyze the experiments and practitioners' commentaries that touch on this phenomenon. Based on these analyses, I introduce evidence to show how the existing findings could be greatly advanced and articulated by the research in this study. In the final section of this chapter, I discuss the limitations of the state-of-the-art perceived usability instruments and give an overview of how the problem has been addressed in this study.

2.1. Modeling Brand in Communication and Marketing Research

Branding is a core component of the marketing profession. It refers to the activities of positioning a brand in the right segments of the consumer market, communicating the related values and promises, and managing the brand in the long term to establish and strengthen its relationship with its target customers (Healey, 2008). In order to effectively conduct these branding activities, the first thing to understand is what a brand actually is and what elements or aspects it encompasses. Aaker (1991) defined a brand as follows:

- A distinguishing name and/or symbol (such as a logo, trademark, or package design) intended to identify the goods or services of either one seller or a group of sellers, and to differentiate those goods or services from those of competitors. -

This definition implies that a brand is embodied as a name or symbol that embraces the rich information that creates a brand's unique identity that customers can recognize or use to distinguish it from other, similar brands. Brands have been studied from various angles and broken up into different sets of sub-constructs. Typically, three of the most well-known brand models are reviewed below.

De Chernatony's (2001) branding model suggested an iterative procedure for creating, implementing, and evaluating a brand. The model first focuses on a strategic, long-term vision of the purpose and values of a brand and the future market in which the brand will need to survive and win. Then, the organization who owns the brand will need to develop a type of working culture congruent with the brand vision. For example, if one of the values of the brand is efficiency (e.g., DHL or FedEx), then, having a slow working style on the part of the staff will bruise the brand's image. Based on the brand vision and the intended organizational culture, marketers should devise actionable branding objectives, analyze the competitive market environment, and define the core attributes of the brand. With the above stages accomplished, the branding plan should be pre-assessed inside the organization, and then launched into the market. The performance of the brand should be evaluated over time.

De Chernatony's model focused on an entire brand management process. Although it also included a final, eight-element sub-model to be used to map a full-fledged brand, it focused more on describing generic branding issues, such as symbol design, risk management, and legal protection. The research in this study used a more analytical model in order to investigate the essential components of a fully-developed brand.

The *brand equity* model presented by Aaker (1991) highlighted the "assets and liabilities" (Aaker, 1991, p. 15) that a brand can bring to the related product or services. These "assets and liabilities" have to do with five brand equity aspects:

- Brand loyalty: Customers become loyal to a brand if they are satisfied with the related product or services. They are likely to buy the same brand repeatedly and are relatively tolerant of moderate price changes. Brand loyalty might be a negative sign for competitors because it generally requires a higher investment in their own brands in order to attract new customers than it does to maintain loyal customers.
- Brand awareness: Before buying a product, customers usually have in mind a set of brands that they want to consider. These brands are known as the "consideration set" (Solomon, 2008, p. 337). Brand awareness focuses on whether a brand has secured a place in the target customers' consideration set and whether it has a good position in the set. Advertising is one of the most effective ways for raising brand awareness (Solomon, 2008).

- Perceived quality: The characteristics of a product or type of service are pre-specified by inventors, managers, or marketers; however, how these characteristics are actually perceived by customers is a different story. Positive perceptions might result in higher purchase intentions and brand loyalty (Aaker, 1991). In contrast, negative perceptions can be quite difficult for a brand to recover from (James, 2010);
- Brand associations: A brand is often depicted by customers through a variety of associations, including its associations with a country, type of personality, or spokesperson. One of the most typical examples is sports clothing brands, such as Nike and Reebok, which usually associate their products with popular sports stars in order to show how well their products can support athletic performances. Establishing appropriate brand associations is an effective way to form intended brand images in the target customers' minds.
- Proprietary brand assets: This aspect focuses on the assets that can help a brand avoid competitor erosion, which includes "patents, trademarks, and channel relationships" (Aaker, 1991, p. 21). Some of these assets can provide legal protection to the names, symbols, or core techniques of an institution.

The brand equity model offers a comprehensive view of all of the properties of a brand. Within the model, this research focused primarily on customers' opinions of product characteristics (perceived quality) and the associations that customers make between the brand and other entities or personality traits (brand associations). In other words, this research investigated the concept of brand perception, which focuses on how a brand is perceived by customers (Berger, Draganska, & Simonson, 2007; Wänke, Herrmann, & Schaffner, 2007).

In order to understand brand perception, we must first get to know which characteristics and brand associations a brand would like customers to perceive. This question can be answered using the brand identity planning model (Aaker, 1996), which consists of the following key sub-constructs:

- Brand as Product: The soul of a brand is the characteristics of the products and services it represents, which include the products' quality, functions, and appearance. For example, Pringles have been characterized as less greasy (or healthier) than

other chips on the market and its stable paper package guarantees that the chips can be kept whole upon delivery to customers.

- **Brand as Organization:** These are the traits that customers associate with the organization that owns the brand. Typically, the organization possesses more long-held characteristics than those associated with the products. For example, Procter & Gamble, who owns the Pringles brand, is known as a large company with a strong R&D capacity and broad coverage of product categories.

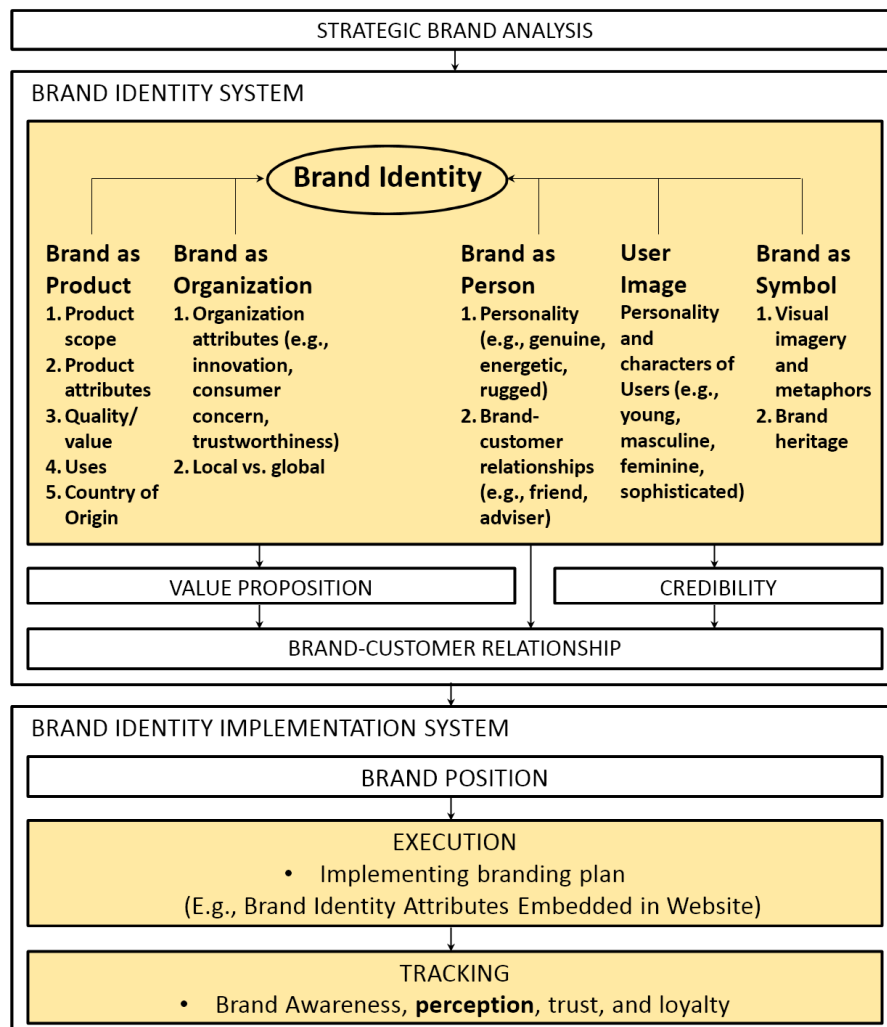


Figure 2. Aaker's (1996) Brand Identity Planning Model.

- **Brand as Person:** Similar to humans, brands also have their own personalities. Aaker (1996) suggested that brand personality can span five general dimensions: “sincerity, excitement, competence, sophistication, and ruggedness” (Aaker, 1996, p 144).

Levi's, for example, can be associated with a tough, durable, and frugal personality. Peoples' tendencies to anthropomorphize brands reflect the need of establishing deep, long-term relationships with the brand, its products and the experience it engenders.

- **User Image:** User image focuses on the typical characteristics of the main user group of a particular brand. In other words, how are the loyal customers of a brand perceived by people? For example, Louis Vuitton is considered to be an upper class brand, owned by those individuals into fashion who have a high social status. Another example can be found in famous slogans intended to shape users' images of particular brands, such as Apple's "Think Different" and Nike's "Just Do It." The user image may or may not be congruent with the brand's personality (Parker, 2009). For example, Levi's user image has been associated with being urban, hip, and contemporary, but not frugal and durable.
- **Brand as Symbol.** A symbol is the visual identifier of a brand. A good symbol is easily recalled or recognized by customers. Known examples include McDonald's big shiny M and Apple's bitten apple.

As indicated by Aaker's model (Figure 2), by analyzing the market environment (e.g., competitors, customers, and the brand itself), marketing professionals are able to define the identity of a brand based on the five constructs reviewed above. Then, the attributes of the brand are refined and enriched by analyzing the potential benefits (value proposition), credibility issues, and expected brand-customer relationship. The fully-developed brand identity system is specialized or adjusted afterward to fit into the particular characteristics of the targeted customer types (brand position). In the end, the finalized branding plan is executed both in terms of designing the actual product and services and through other promotional and communication channels, such as advertising, event sponsorship, and the web. The performance of the brand can be tracked and evaluated in regard to the awareness, perception, trustworthiness, and customer loyalty of the brand.

In this research study, the brand identity planning model was used to describe the various attributes of brand perception. This model is more analytical than De Chernatony's procedural model in terms of formulating a fully-developed brand and more focused than the brand equity model.

2.2. Online Branding Strategies and Frameworks

Most of the exiting eBranding frameworks are put forward by marketing professionals (De Chernatony, 2001; De Chernatony & Christodoulides, 2004; Page & Lepkowska-White, 2002; Ries & Ries, 2000; Simmons, 2007). First, these frameworks try to differentiate eBranding from bricks and mortar branding. Ries and Ries (2000) suggested that one of the main differences between eBranding and bricks and mortar branding can be described as “the law of singularity” (Ries and Ries, 2000, p. 73). That is, in bricks and mortar stores, the retailers might carry at least two competing brands as leverage for their businesses to succeed. For example, if a supermarket manager cannot persuade Coca-Cola to participate in a weekly promotion, s/he can always resort to asking Pepsi (Ries & Ries, 2000). In contrast, on the Internet, a manufacturer could sell its product directly to customers without having to go through a retailer. Therefore, less dominant brands might completely lose their chances. For example, customers might always use their favorite ecommerce website (e.g., Amazon) without bothering to consider a secondary choice. Although this law is not true for every product category and was more applicable to the less mature, online market that existed a decade ago, instead of the one that exists today, it illustrates the more intimate relationship between a brand and its customers on the Internet than in real life – transactions are not necessarily mediated by retailers any more.

Another important characteristic of eBranding is interactivity (De Chernatony, 2001; Ries & Ries, 2000). Some offline branding methods, such as newspaper advertisements and TV commercials, put customers into a more passive role. Customers mostly just read or watch what is presented to them without needing to respond to or take control over anything. Websites, in contrast, give users the freedom to decide where they want to go and what they want to see in a complex information structure (De Chernatony, 2001; Solomon, 2008). They allow users to “feel, taste, and smell” a brand without actually going to a bricks and mortar store (Solomon, 2008, p. 7). This point explains, from a branding perspective, the importance of making a website easy-to-use.

Besides closer brand-customer relationship and interactivity, De Chernatony’ eBranding framework also stresses the need to build online customer communities because web users value their capacity to exchange thoughts with other users (De Chernatony & Christodoulides, 2004). Simmons (2007) highlighted the importance for a company to

understand the moral, ethical, and emotional values that are highly regarded by customers and the need for the company to convey these values through its website. Page and Lepkowska-White (2002) showed that the communication of customer-recognized values can be greatly influenced by four web design attributes: accessibility, navigation, quality of content, and customizability.

Existing eBranding frameworks are devoted to answering the questions of *what to do* and *how to do it* in order to devise strong eBranding plans and strategies. These frameworks fail to take into consideration, however, the question of *how well* the plans and strategies can help effectively communicate the intended identity of a brand. More specifically, how well an organization's website can communicate the organization's identity traits? Does the website have a positive or negative impact on the image of the organization? If it is negative, how can the website be improved or redesigned? In order to answer these questions, an eBranding evaluation method should be developed in order to systematically assess the branding effectiveness of an existing organizational website and provide practical recommendations for improvement (Chapter 3).

2.3. Reciprocal Effects of Website Visiting and Brand Perception

Research has been conducted to study how engagement with a known brand affects a user's judgment of a website and its usability (De Angeli, Hartmann, & Sutcliffe, 2009). Users with positive opinions toward a brand give significantly better usability and aesthetics ratings than those who hold negative opinions.

This research (Chapter 5) investigates an opposite issue, which may have a broader impact: how the quality of design may affect the perception of the brand. Existing studies explore this issue from a relatively holistic perspective. That is, either website design quality or brand perception is considered a single-dimensional construct. For example, an overall positive attitude toward a website could make users attach more positive personality traits to a brand, such as sincerity (Müller & Chandon, 2003) and innovativeness (Müller & Chandon, 2004). It is unclear, however, which aspect of the design (e.g., content, graphics, or navigation) gives users a more sincere or innovative feeling. Similarly, using a survey of the users of an online book store, Da Silva and Alwi (2008) found that "corporate brand image" is significantly affected by four qualities of a website: "ease of use, security, personalization, and customer care" (Da Silva and Alwi,

2008, p. 1052). Neither corporate brand image nor ease of use is broken down into more specific aspects. Lepp, Gibson, and Lane (2011) demonstrated a successful case in which prospective visitors' perceived risk of a tourist destination (i.e., Uganda) was significantly reduced after visiting the tourist destination's official website. Capitalizing on the study results, the authors proposed a five-dimensional model to help reducing tourists' perceived risk. The model, however, focused more on the causes of the problematic images of African countries (e.g., cultural barriers) than on website design. In addition, Macias (2003) investigated the importance of interactivity on the persuasiveness of online advertisements. However, *interactivity* is also a relatively high-level construct.

In addition to lacking a higher level granularity, most related works utilize existing brands and websites as test beds. A key limitation of this approach, however, is the lack of control over users' prior brand knowledge. In particular, Lee, Hong, and Lee (2004) suggested that people who are highly knowledgeable about a brand are unlikely to associate its image to the quality of the brand's website. Jansen et al. (2007; 2009) indicate that the reputation of a search engine can affect users' perceived quality of the search results: users give higher ratings to the search results from a famous than from a less-known search engine, although the search results are the same (Jansen, Zhang, & Schultz, 2009; Jansen, Zhang, & Zhang, 2007). Therefore, in order to see the *pure* effects of website design quality on brand perceptions, users' prior brand knowledge must be strictly controlled. In Chapter 5, I used a fictitious, generic brand name to minimize the impact of prior brand knowledge.

2.4. Usability for Branding

HCI practitioners are aware of the relevance of branding to the interaction design field. Marcus (2011) suggested that "user experience professionals can benefit from learning more about branding's role in managing customer experience" (Marcus, 2011, p. 32). He also highlighted the usefulness of establishing branding models from interaction designers' perspective: "[these models] might be especially useful to help the CHI community understand the substance and benefit of branding" (Marcus, 2004, p. 19). In terms of eBranding, HCI practitioners particularly stress the role of website usability. Spool (1996; 2002) suggested that, most of the time, users' online behaviors are goal-oriented. Their perceptions of a brand are more determined by whether the brand's

website can help them achieve their goals than whether the website is aesthetically pleasing. Baty (2006) expressed similar thoughts and found that errors that occurred during web browsing might have a damaging impact on brand image. McGovern (2001) highlighted the importance of website content and showed that users' main online activities were focused on content gathering. Therefore, content was the main vehicle for conveying branding messages. In addition, given the exponential growth of online information, content should be well written in order to attract users.

As shown above, the impact of website usability on branding has been often raised by practitioners in the past two decades. Surprisingly, this important issue has received limited empirical scrutiny to date. For example, a previous research study has addressed the relationship between usability as a whole and brand values (Bolchini, Garzotto, & Sorce, 2009), with the known limit of not providing analytical guidance on which factors of usability most affect brand perception. Subsequent work (Garzotto et al., 2010) studied aesthetic attributes as elements of perceived usability and their effect on brand perception. Key limitations remain as to the lack of control over prior brand knowledge and the confounding variables derived from a variety of website domains investigated.

2.5. Methods for Capturing Perceived Usability and Their Limitations

The ISO's (International Organization for Standardization) definition of usability is "the effectiveness, efficiency and satisfaction with which specified users can achieve specified goals in particular environments" (ISO 9241-11, 1998). From the users' perspectives, how they judge the usability of a product is described as *perceived usability*. McGee, Rich, and Dumas (2004, p. 909) suggested that perceived usability is "the users' perception of how consistent, efficient, productive, organized, easy to use, intuitive, and straightforward it is to accomplish tasks within a system." One of the most popular ways to measure perceived usability is to ask users to complete a questionnaire through which they can provide their own ratings toward predefined usability aspects of a product.

Through a survey on the existing literature on perceived usability questionnaire development, I found that three main types of perceived usability questionnaires exist: universal perceived usability questionnaires, perceived usability questionnaires for websites, and perceived usability questionnaires for mobile applications. In the following

sections, I will comprehensively review the typical questionnaires in those three categories and, then, by comparing them with heuristic usability evaluation (analytical usability), I will suggest one of the major limitations that the existing perceived usability questionnaires have.

2.5.1. Universal Perceived Usability Questionnaires

Some of the perceived usability questionnaires could be used to evaluate the usability of any type of electronic product. Considering usability as a generic concept, these questionnaires focus on the most common usability experiences.

The Practical Heuristics for Usability Evaluation (PHUE) questionnaire was developed by Perlman (1994) and decomposed the concept of perceived usability into three sub-concepts: learning, adapting to the user, and feedback and errors. These sub-concepts were adapted from Nielsen's (1993) 10 heuristics and Norman's (1990) seven principles. On average, four measurement items were used to capture each sub-concept. Each item was assessed on a 7-point semantic differential scale (from Bad to Good) and "Not Applicable" in case an item was not relevant to a specific product. However, this questionnaire has not been validated using standard psychometric methods and is more suitable for assisting usability experts to conduct heuristics inspection rather than to assess perceived usability.

The USE questionnaire (Lund, 2001) operationalized perceived usability into four dimensions through an exploratory factor analysis: usefulness, ease of use, ease of learning, and satisfaction. The questionnaire items were either elicited from existing literature or created through brainstorming. Seven-point Likert scales (from Strongly Disagree to Strongly Agree, and Not Applicable) were used to measure these items. The questionnaire was empirically validated by asking users to evaluate the usability of a number of products. However, the validity of the questionnaire has not been double-checked through a confirmatory factor analysis, which is also a critical stage for developing standardized questionnaires.

The QUIS (Questionnaire for User Interface Satisfaction, Chin et al., 1988) captured the perceived usability of information systems from five aspects: overall reaction to the software, screen, terminology and system information, learning, and system capabilities.

This questionnaire is highly reliable and valid because it has already been tested and validated many times by applying it to the evaluation of different types of products (Chin et al., 1988). Various semantic differential scales are used to assess the items in this questionnaire. For example, the item “reading characters on the screen” is measured on a 9-point semantic differential scale from Hard to Easy. However, the QUIS questionnaire was developed 20 years ago; therefore, its reliability and validity need to be retested with state-of-the-art digital products.

The Perceived Usefulness and Ease of Use (Davis, 1989) questionnaire was developed based on the assumption that how much users would like to use an information system depends upon how much the system is useful and easy to use. Then, according to the definitions of “usefulness” and “ease of use” in the existing literature, candidate questionnaire items were created and tested either with finished electronic products or prototypes (e.g., to test a design concept would be useful and easy to use in the future). This questionnaire is more suitable for measuring the overall user experience than detecting specific usability problems.

Computer Usability Satisfaction Questionnaires (Lewis, 1995) contain a set of four questionnaires developed by IBM. Most of the items in these questionnaires were created by usability professionals based on their experiences developing and evaluating computational systems. These questionnaires were designed to serve different purposes: the After-Scenario Questionnaire and Printer-Scenario Questionnaire were used to capture users’ feelings toward performing scenario-based tasks on an information system, whereas the Post-Study System Usability Questionnaire and Computer System Usability Questionnaire (CSUQ) were used for evaluating users’ “overall satisfaction” (Lewis, 1995), including system usefulness, information quality, and interface quality. However, one limitation of the Post-Study System Usability Questionnaire and CSUQ is that the scope of the “interface quality” is broad, whereas only three items were used to measure this construct. This is not enough to reflect the complexity of the interface design of a computational system.

The SUS (System Usability Scale) is a short questionnaire (only 10 questions) developed by Brooke (1996) and aimed at quickly and reliably assessing the usability of information systems. Brooke suggested that the results of usability evaluations vary

significantly according to the *context* of where a system is used, such as within a selected user group, task scenarios, or social context. Therefore, instead of developing a lengthy questionnaire trying to cover every circumstance of usability, what practitioners truly needed was a simple, but reliable tool that could capture the essence of the usability of a system in a brief time period. Based on this idea, only 10 well-refined and extensively tested items were created for the SUS. As purposed by Brooke, the SUS was more useful for practitioners who expected quick results than researchers who wanted to get a comprehensive view of system usability. In order to further shorten the SUS scale and align the questionnaire items with the ISO (1998) definition, the Usability Metric for User Experience (Finstad, 2010) was developed, which is a four-item questionnaire that measures effectiveness, satisfaction, efficiency, and overall usability.

The Purdue Usability Testing Questionnaire (PUTQ; Lin, Choon, & Salvendy, 1997) was developed based on the theory of human information processing. By considering human-computer interaction as a process in which users perceive and process the information delivered by the interface of an information system, Lin et al. (1997) elicited eight critical factors: compatibility, consistency, flexibility, learnability, minimal action, minimal memory load, perceptual limitation, and user guidance. In order to measure these factors, 100 measurement items were created based on the existing usability guidelines and principles. The advantage for the PUTQ is that it offers an exhaustive list of usability issues for information systems. The disadvantage, however, is that it is difficult for users to stay focused when attempting complete the entire questionnaire.

The questionnaires reviewed above are applicable to the evaluation of any type of electronic product. However, most of them are too general to be suitable for evaluating domain-specific usability dimensions. For example, most of these questionnaires do not capture the concept of navigation, which is critical for website usability. Although the PUTQ provides a list of 100 items that cover almost every aspect of any product, it takes too long for users to finish. As a solution to this problem, domain-specific perceived usability questionnaires were developed.

2.5.2. Perceived Usability Questionnaires for Websites

WAMMI (Website Analysis and Measurement Inventory) is one of the most popular services for websites evaluation. It not only provides a questionnaire, but also a large

database of analytical evaluation methods. Capitalizing on these tools and resources, it can generate a systematic evaluation report for a specific client's website. The WAMMI questionnaire covers five "ease of use" factors (Kirakowski, Claridge, & Whitehand, 1998): attractiveness, control, efficiency, helpfulness, and learnability. In order to measure each of these factors, candidate measurement items were created by interviewing website professionals and users. Then, the candidate items were applied to the evaluation of a number of real-world websites for extensive testing and validation (Kirakowski, Claridge, & Whitehand, 1998).

Wang and Senecal's (2008) perceived website usability questionnaire focused on three usability dimensions: ease-of-navigation, speed, and interactivity. In addition, noticing that "perceived disorientation" (i.e., getting lost in a website) was a different construct from ease of use, Ahuja and Webster (2001) created ten measurement items in an attempt to capture the two concepts separately. The items were created by brainstorming with graduate students who were majoring in information systems and surveying existing literature that investigated perceived disorientation.

2.5.3. Perceived Usability Questionnaires for Mobile Applications

Ryu and colleagues (2005; 2006; 2007_a; 2007_b) conducted a series of studies in order to develop a Mobile Phone Usability Questionnaire. Based on the current standards and research on mobile phone usability, the preliminary questionnaire was developed by collecting relevant measurement items from existing usability questionnaires (Ryu & Smith-Jackson, 2005). Then, the questionnaire was further tested and standardized using a larger sample using statistical and analytical methods (for further details, see Ryu and Smith-Jackson (2006)). In addition, their most recent studies showed that, among the six core dimensions of their questionnaire, the *efficiency and control* dimension most influenced the users' decisions in regard to evaluating the usability of a mobile phone (Ryu et al., 2007_a, Ryu et al., 2007_b).

2.5.4. Evaluation-Design Alignment Gap

In order to capture the concept of perceived website usability, I began by looking into the existing perceived usability questionnaires reviewed above. However, I found that few questionnaires were specifically designed for assessing websites and those that were designed that way were either unavailable to the public (WAMMI) or did not

comprehensively cover all the possible aspects of a website (e.g., Wang and Senecal did not assess website content). More importantly, by comparing these perceived usability questionnaires with analytical usability evaluation methods (e.g., heuristic evaluation), I found that these questionnaires were limited in terms of the interpretability of the evaluation results and the quickness to inform redesign or improvement strategies. Yang, Linder, and Bolchini (2012) defined this problem as an Evaluation-Design Alignment Gap.

Let us consider the following scenario. A web designer receives a usability testing report in which the following finding is reported in the results section of a perceived usability questionnaire: 80% of the participants strongly disagreed that the interaction with the system is clear and understandable. This type of finding clearly diagnoses a problematic aspect of the user experience with the system. It fails, however, to inform the evaluators and designers about one fundamental concern, which can be summarized by the questions “What should we do with this finding?” and “How can this finding lead us to consider new design requirements?”

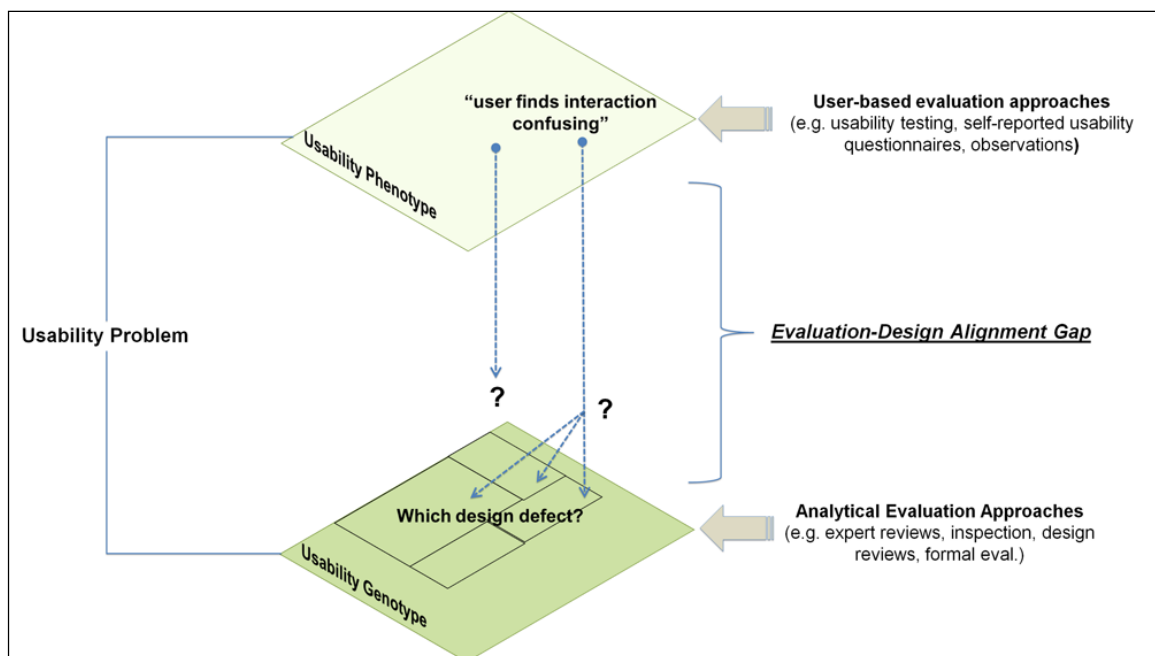


Figure 3. Addressing the Evaluation-Design Alignment Gap can help relate usability problems to the analytical composition of the system design.

This scenario is a good illustration of the Evaluation-Design Alignment Gap. In simple terms, the Evaluation-Design Alignment is the property of a usability instrument to gradually drive the evaluation activity toward requirements for redesign. The more the evaluation feedback is generic and not conducive to recommend actionable design requirements, the poorer the Evaluation-Design Alignment. In the example just noted, which components of the design are perceived as not clear and understandable? Is it the system of labels, or is it the overall navigation architecture? Is it the content presented? Is it the graphical layout of the information? Once these aspects are clarified, we then notice that the notion of clear and understandable means different things when applied to different design elements: labels that are perceived as not clear may be ambiguous, vague, redundant, or too technical. Navigation that is perceived as unclear may provide poor orientation or excessively deep hierarchies. Unclear graphical presentations may indicate issues concerned with visual ordering, consistent layout grids, alignment and spacing of the elements, and so on.

A usability problem, defined as an obstacle to a satisfactory, efficient, and effective use of the system, invariably features two components: the perceived manifestation of the obstacle on the actual user experience (i.e., usability phenotype) and the actual defect(s) of system design that is causing the problem (i.e., usability genotype) (Lavery, Cockton, & Atkinson, 1997; see Figure 3). Whereas analytical evaluation methods (e.g., expert reviews, heuristic inspections, and formal evaluations) tend to unveil the nature of the system design defects (i.e., the space of the usability genotype), user-based evaluations (e.g., usability testing, perceived usability questionnaires) tend to capture what problems users perceive or encounter in their experiences with the system.

Usability problems—as captured by existing perceived usability questionnaires—emerge at the level of the usability phenotype. Eventually, however, evaluators need to identify the cause of this problem in terms of design defects. The target space to be investigated then moves from the usability phenotype to the genotype (i.e., the set of analytical design components that constitute the system under evaluation). What instruments can be developed to bridge this Evaluation-Design Alignment Gap? How can evaluators and designers be supported in seamlessly and systematically relating the user-reported findings into actionable design drivers? These questions are addressed in Chapter 4 – DEEP: DDesign-oriented Evaluation of Perceived website usability.

In the following chapter, I present an evaluation approach for systematically evaluating the branding effectiveness of content-intensive websites. This approach attempts to address the research problem discussed in section 2.2 (i.e., existing eBranding frameworks do not cover systematic evaluation methods).

Chapter 3. BREW: A Systematic Approach for Evaluating the Branding Effectiveness of Content-intensive Websites

A systematic evaluation approach is a necessary supplement for existing online branding methods (Chapter 2, Section 2.2). It allows eBranding professionals to evaluate the branding effectiveness of their current web applications and make redesign or improvement strategies accordingly. In previously published works (Bolchini, Yang, & Garzotto, 2009; Yang & Bolchini, 2010), a preliminary evaluation framework was developed in order to explore the feasibility of devising such an approach. This framework proposed a step-by-step guide by which to prepare the related eBranding evaluation instruments (e.g., task scenarios and questionnaires), conduct the evaluation with targeted users, and analyze the results. More importantly, the framework showed a high potential in regard to supporting the intended eBranding evaluation purposes when it applied to the evaluation of two real-world branded websites (Bolchini, Yang, & Garzotto, 2009). In spite of these promising findings, however, this framework has two major limitations:

- 1) The framework is not supported by a robust branding theory. The method that it used to analyze the key attributes of a brand is not systematic enough. For example, it does not take into account user imagery (Aaker, 1996).
- 2) The framework does not provide enough guidance to interpret the evaluation results. It is up to an eBranding researcher to decide whether the score that a brand attribute received is high or low, and no illustration is provided in regard to how the evaluation results can be used to improve website design.

In this chapter, I propose an enhanced version of this evaluation framework by solidifying it with Aaker's brand identity planning model (Aaker, 1996) and substantially strengthening the results interpretation section. The name of the improved evaluation approach is BREW: A systematic approach for evaluating the BBranding Effectiveness of content-intensive Websites.

3.1. The BREW eBranding Evaluation Approach

The goal of the BREW evaluation approach is to evaluate how well a company or an institution's website can communicate the company or institution's key branding attributes. This approach can be used by any company or institution that wants to check whether any of their branding attributes are not effectively communicated by their websites and come up with practical improvement strategies. The BREW approach consists of five phases: brand attribute inspection, brand perception questionnaire composition, task design, user test, and results analysis (Figure 4). In order to illustrate how each phase can be conducted, the BREW approach was applied to a case study on evaluating of two branded websites: AT&T (www.att.com) and Verizon (www.verizon.com). The reason why I used two branded websites (instead of only using one) is explained below.

- As suggested by existing branding evaluation methods, the most effective way to assess the performance of a brand is to compare it to its competitors (Ambler, 2003; De Chernatony, 2006; De Chernatony & McDonald, 2003). The aim of a brand is to perform equally to or better than its competitors, but not necessarily strive to achieve optimal status. For example, if most of the mid-level pc laptops on the market have an Intel® Core™ i3 processor, a pc laptop in the same price range may win with an Intel® Core™ i5 processor, but does not need to have an i7 processor to win in the market. BREW partly adopted a similar strategy: if the average score that a brand received on a particular attribute is equal to or better than that of its competitors, it is considered that the attribute is satisfactorily perceived. For example, suppose we evaluate the *sincerity* of a brand on a 10-point scale. If the main competitor of the brand received a sincerity rating of 6/10, then the brand may only need an equal or slightly higher rating (e.g., 6.5/10) to stay competitive in the market. It is unnecessary, at least in a certain period, for the brand to make substantial improvements to try to achieve the highest possible score (i.e., 10/10). Therefore, to use the BREW approach to evaluate the branding effectiveness of a website, only knowing the evaluation results (brand perception ratings) of the website under evaluation is not enough. In order to interpret the results, an eBranding researcher should at least know how well a main competitor's website performs in communicating similar brand attributes.

- In the case study, AT&T was taken as the website under evaluation. Verizon was chosen among the competitors of AT&T (Yahoo! Finance, 2013) because it has the most similar market cap and revenue to AT&T. In this way, the evaluation results of the two branded websites were more comparable, and it made more sense for the research to interpret the branding effectiveness of the AT&T website when compared to the Verizon website.

It is worth noting that the main aim of the case study is to showcase the BREW approach in action. Some of the materials used may not truly reflect the actual branding objectives of the company. Real-world evaluation projects should be conducted by a company's web user experience and marketing team, as they have a more thorough and accurate understanding of their own brand and website.

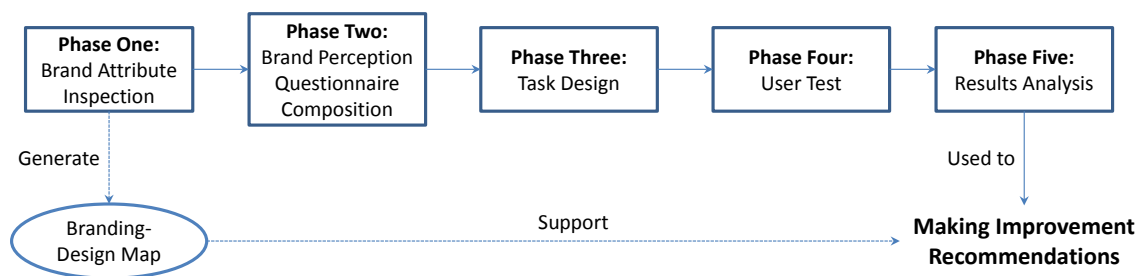


Figure 4. An overview of the BREW evaluation approach.

3.1.1. Phase One: Brand Attribute Inspection

Before starting the evaluation, the marketing team of the organization who owns the brand needs to clearly define the intended *identity of the brand* (i.e., the perceptual objective of the branding practice). The creation of this definition could be done with the help of Aaker's brand identity planning model (Aaker, 1996). The marketing team members could ask themselves the following questions to classify the brand's identity attributes into four categories:

- What are the key product characteristics that the brand wants its customers to know? (Brand as Product).
- Does the organization or company that owns the brand have any long-held beliefs? If so, what are those beliefs? (Brand as Organization)
- What are the characteristics of the target users/customers of the brand? (User Image)

- If we consider this brand to be a person, how does s/he want other people to think of him/her? (Brand as Person)

The BREW evaluation approach does not cover the “Brand as Symbol” dimension (Aaker, 1996) because that dimension has more to do with the graphical design of the symbol of a brand. The effectiveness of a symbol is unlikely to be affected by how a website is designed, as long as the symbol is placed in a visible area on the website interfaces.

In order to illustrate the BREW evaluation, the case study elicited AT&T’s brand identity attributes from its published materials, including its online company profile, value statement (AT&T_a, 2013), and recent TV commercials. These materials are either published on AT&T’s official website or its YouTube channel (AT&T_b, 2013). I read and/or watched these materials, identified the key brand attributes, and categorized them into Aaker’s system. The complete brand identity system for AT&T is shown in Table 1. A total of 36 key brand identity attributes were collected. Again, I used these attributes to show how BREW works. In real-world scenario, the brand identity system should be developed by the marketing team of the company.

Table 1. Eliciting AT&T’s key brand identity attributes.

Brand as Product	Brand as Organization	Brand as Person	User Image
1. Innovative Products	12. Committed to Innovation	20. Valuable	32. Fast
2. Reliable Products	13. Understands What Its Customers Want	21. Creative	33. Ambitious
3. High-Quality Products	14. Delivers What Its Customers Want	22. Caring	34. Confident
4. Innovative Services	15. Makes Things Easy For Its Customers	23. Adventurous	35. Saving
5. Reliable Services	16. Empowers Its Customers To Use Technology Safely	24. Righteous	36. Young
6. High-Quality Services	17. Empower Its Customers To Use Technology Securely	25. Reliable	
7. Highly International Services	18. Committed To Do the Right Thing	26. Secure	
8. Has the Nation's Largest 4G Network	19. Connects Their Users To the World In Which They Live and Work	27. Successful	
9. Operates the Nation's Largest Wi-Fi Network		28. Leader	
10. Provides a Better TV Experience		29. Supportive	
11. Provides Excellent Customer Care		30. Understanding	
		31. Intelligent	

The main practice of Phase One is to conduct a *brand attribute inspection* of the website under evaluation to see whether and how a brand attribute is conveyed by it. The inspection can be conducted using the following steps:

- 1) Find or create a sitemap for the branded website under evaluation.
- 2) Create a new spreadsheet and list all of the brand identity attributes in the left column.
- 3) Insert several new rows at the top of the spreadsheet (Figure 5).
- 4) Starting from the first branch of the sitemap, list the titles of the related webpages horizontally at the top of the spreadsheet and clearly mark to which branch each page belongs.
- 5) Inspect the content and design of each page in order to determine whether it can communicate one or several of the brand identity attributes. Check (“X”) all relevant attributes on the sheet. The following two criteria can be used for deciding whether a brand identity attribute is conveyed by a page.
 - a. The brand attribute or its synonym is directly communicated through the text, pictures, or embedded video clips. For example, on a number of pages within the website, AT&T explicitly mentioned that its products have high *reliability* (reliable) and “come at great *value*” (valuable).
 - b. The brand attribute is not directly conveyed in verbatim or in synonym, but implicitly communicated using analogies, metaphors, or non-verbal cues (e.g., via the color scheme or specific design features). The brand attribute should be easily perceived by the users even if it is implicit. For example, the pop-up live chat on almost every page of the AT&T website implies one aspect of *excellent customer care*. Also, the special offers (e.g., the Mother’s Day offers) on AT&T’s homepage may suggest that AT&T *understands and delivers what its customers want*.

The steps above can also be applied to a selected set of branches if the entire site is not going to be evaluated. In order to enhance the inter-rater reliability, the inspection should be conducted separately by at least three eBranding experts. Then, they should conduct a consolidation session in order to compare and consolidate the inspection results. The final inspection map must be agreed upon all three experts.

In the case study, four branches of the AT&T website were inspected: Digital TV, High Speed Internet, Home Phone, and Wireless Cell Phones & Devices. A snippet of the inspection results is shown in Figure 5.

		Digital TV			AT&T U-verse Total Home DVR	AT&T DIRECTV	High Speed Internet		AT&T Wi-Fi
		Homepage	U-verse TV	AT&T U-verse Online			U-verse High Speed Internet		
3	Brand as Product								
4	Innovative Products	X	X		X				
5	Reliable Products		X		X				
6	High-Quality Products	X	X		X				
7	Innovative Services	X	X			X			
8	Reliable Services		X				X		
9	High-Quality Services		X	X		X	X		X
10	Highly International Services								
11	Has the nation's largest 4G network	X							
12	Operates the nation's largest Wi-Fi network								
13	Provides a better TV experience		X	X					
14	Provides excellent customer care	X	X	X					
15									
16	Brand as Organization								
17	Committed to innovation		X		X	X			
18	Understands what its customers want	X	X	X	X	X	X		X
19	Delivers what its customers want	X	X	X	X	X	X		X
20	Makes things easy for its customers	X	X	X	X	X	X		X
21	Empowers its customers to use technology safely		X				X		
22	Empower its customers to use technology securely		X				X		
23	Committed to do the right thing	X	X				X		
24	Connects their users to the world in which they live and work	X	X				X		X
25									
26	Brand as Person								
27	Valuable	X	X	X	X	X	X		X
28	Creative		X		X	X			
29	Caring	X	X	X	X	X	X		X
30	Adventurous		X						
31	Righteous						X		
32	Reliable		X		X		X		
33	Secure		X				X		

Figure 5. Inspecting where the branding identity attributes are communicated within the website.

The significance of this inspection is two-fold. First, it offers a round of expert-based evaluation (similar to a heuristic evaluation). An eBranding expert can directly point out which brand identity attributes are not embedded in or effectively communicated by the website. These findings can serve as a complement to the user test in Phase Four. For example, as highlighted in Figure 5, two of the brand identity attributes, *highly international services* and *righteous*, are not embedded in most of the webpages, which may weaken the users' perceptions. Second, this inspection reveals the locations of the brand identity attributes in the website, which not only illustrates the intensity of a brand identity attribute being communicated by the website, but also provides a map by which to interpret the users' test results in Phase Five. For example, in case the attribute *valuable* is not well-perceived by users (Figure 5), researchers may follow the map to check the actual pages where the attribute is embedded in an attempt to discover why these pages did not communicate the attribute well.

3.1.2. Phase Two: Brand Perception Questionnaire Composition

The main objective of this phase is to transform the elicited brand identity attributes into a brand perception questionnaire. This questionnaire will be used in Phase Four to capture users' opinions toward a brand after using the brand's website. Since it is possible that users will not know a certain aspect about a brand, the questionnaire can start with the following instruction:

Based on your experiences with the website, please fill out the following questionnaire regarding your opinions toward [name of the brand]. It is possible that you will not know every item listed below. If that is the case, please provide a reasonable guess according to your impression of the website.

Then, the brand identity attributes related to *brand as product* and *brand as organization* can be transformed into short statements with each statement measured by a 5- or 7-point Likert scale that ranges from *strongly disagree* to *strongly agree*. For example, the brand identity attribute *innovative products* can be transformed into the following questionnaire item:

The products that AT&T provides are innovative.

Strongly Disagree ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Strongly Agree

The attributes related to *user image* do not focus on the brand itself, but on the characteristics of the brand's users. Therefore, before asking the test participants to rate the user image attributes, a clear instruction, such as that below, should be given:

Based on your experiences with the website, what do you think of the people who buy and use [name of the brand]? Please provide your ratings for the following items.

Each user image attribute can also be transformed into a statement and measured by a Likert scale. For example,

AT&T users are saving money.

Strongly Disagree ----- 1 ----- 2 ----- 3 ----- 4 (I Don't Know/Neutral) ----- 5 ----- 6 ----- 7 ----- Strongly Agree

However, it is sometimes also interesting to know whether participants think that the opposite side of a character is true. For example, in case people disagree that AT&T users are saving money, it is important to ask if they think that the AT&T users tend to *waste* money. In this case, a semantic differential scale (instead of using two Likert scale items) can be used to measure two opposite user image attributes.

Wasting ----- 1 ----- 2 ----- 3 ----- 4 (I Don't Know/Neutral) ----- 5 ----- 6 ----- 7 ----- Saving

The attributes relating to *brand as person* need the test users to anthropomorphize the brand under evaluation. The following instruction can be given before asking the users to rate the brand as person attributes:

*If you consider [name of the brand] as a person, what is your opinion toward him/her?
Please rate the following items.*

Similar to the user image attributes, each brand as person attribute can also be measured using either a Likert or semantic differential scale. The full brand perception questionnaire used in the case study can be found in Appendix A.

3.1.3. Phase Three: Task Design

The aim of this phase is to design a set of tasks to be used to motivate users to explore and experience the website. The suggested steps to designing user tasks follow the goal-oriented approach (Bolchini & Mylopoulos, 2003), which focuses on analyzing the profiles of the targeted users and the specific goals they want to achieve.

- 1) Based on the content and functionalities of the website sections on which the brand attribute inspection was performed, clearly define the typical user profiles of these website sections. For example, the four sections of the AT&T website (i.e., Digital TV, High Speed Internet, Home Phone, and Wireless Cell Phones & Devices) correspond to four types of users: TV users, Internet users, home phone users, and cell phone users.
- 2) Write usage scenarios for each type of user in order to clarify the specific goals that they want to achieve by using the website. The following is an example of the scenarios written for TV users.

Joan was interested in receiving digital TV in her home. After talking to her friends, she realized that the U-verse service offered by AT&T might be a good option. Thus, she went to the AT&T website to look for more details. While reading about the prices and key features of the various U-verse plans, an idea popped to her mind: "It

would be cool if I could record some of my favorite TV shows.” Therefore, she checked the functionalities of the U-verse DVR receiver. She found that the receiver was able to store up to 65 hours of HD content, which was more than enough to store several full episodes of her favorite shows.

- 3) Pick the most critical user goals and rephrase them into meaningful user tasks. For example, the scenario above can be rephrased into the following task.

Task: Suppose you are interested in receiving digital TV in your home and want to record some of your favorite TV shows. Please check how many hours of HD video content you can record with the U-verse DVR.

In order to ensure that the users extensively experience the website, three to eight user tasks should be designed. A full list of tasks used in the case study can be found in Appendix B.

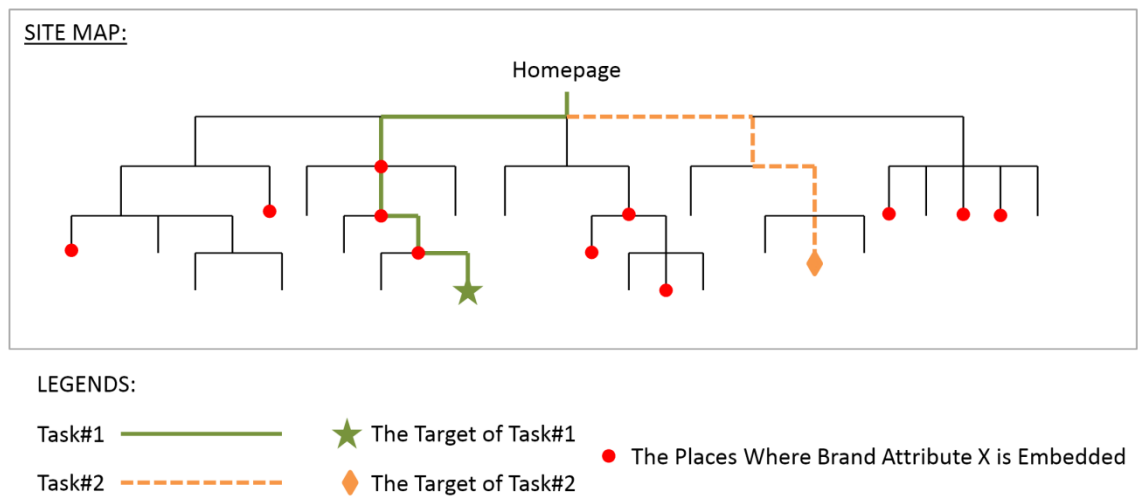


Figure 6. Realistic tasks allow researchers to examine whether users can naturally perceive the intended brand identity attributes.

Tasks designed in the above manner allow eBranding researchers to investigate whether a website can naturally communicate the intended brand identity attributes to the users in the process of accomplishing their typical tasks. If the brand attributes are effectively communicated along the path of accomplishing a task (Figure 6, Task#1), a user may have a better chance at perceiving them. On the contrary, if the brand

attributes are embedded in webpages not relevant to typical tasks, then the chance for the users to see and perceive the attributes may be much lower (Figure 6, Task#2). This also warns eBranding researchers to avoid designing *leading* tasks (e.g., tasks that intentionally lead users to places where brand attributes are communicated). When designing user tasks, researchers ignore where the brand attributes are located and, instead, focus on the user goals.

3.1.4. Phase Four: User Test

In this phase, the eBranding researchers need to devise a procedure and strategies by which to conduct a user test. The aim is to see whether a brand's website can improve or enhance users' satisfactions with the brand. The website should, at least, not negatively affect the users' prior knowledge about the brand. More importantly, researchers can use this test to learn whether the brand's website is in a competitive position in terms of branding effectiveness when compared to its competitors' websites. This phase can be conducted using the following steps.

- 1) Before starting the user test, one critical factor that needs to be captured is the users' prior knowledge of the brand. Users may gain knowledge of a brand through many channels, such as watching TV commercials, buying and using related products, and browsing the Internet. Previous studies have shown that the more a user is knowledgeable about a brand, the less likely his/her opinions toward the brand can be affected by the brand's website (Lee, Hong, & Lee, 2004). Therefore, when analyzing the evaluation results, researchers should divide users into different groups according to their prior brand knowledge and investigate the branding effectiveness of the website separately for each group. In order to capture a user's prior brand knowledge, a pre-test questionnaire can be created using the following two items.

Item#1. Are you familiar with [name of the brand]?

Never Heard of It ----- 1 ----- 2 ----- 3 ----- 4 ----- 5 ----- 6 ----- 7 ----- Very Familiar

Item#2. What is your overall impression towards [name of the brand]?

Very Negative ----- 1 ----- 2 ----- 3 ----- 4 (I Don't Know/Neutral) ----- 5 ----- 6 ----- 7 ----- Very Positive

The first item asks the users whether and how much they know about the brand, while the second item captures the direction (positive/negative) of the users' opinions. The reason for having the second item is to examine whether the users' opinions can be changed by visiting the brand's website. This can be done by asking the users to fill in the brand perception questionnaire (composed in Phase Two) after performing tasks on the website. Then, the researchers would need to compare the rating of each brand identity attribute with the rating for Item#2. In order to make this comparison easier, researchers may use the same rating scales (5-point/7-point) as the ones used in the brand perception questionnaire.

A more precise way to capture prior brand knowledge is to ask the users to fill in the brand perception questionnaire both before and after the test. This approach allows researchers to see the change in each brand attribute before and after website visiting. This approach, however, may sensitize users in the pre-test, which encourages them to give different ratings in the post-test (Churchill & Iacobucci, 2009). In order to avoid this issue, instead of capturing prior brand knowledge with the testing group (i.e., the group of users who will be asked to perform tasks on the website), researchers may recruit a control group, which is made of the same type of users as the testing group, to only fill in the brand perception questionnaire (Figure 7). In this way, the impact of the website can be captured by comparing the testing group with the control group. The challenge of this approach, however, is to recruit two groups of users with homogeneous backgrounds (i.e., gender, age, education), something that usually can only be achieved by recruiting large-size samples.

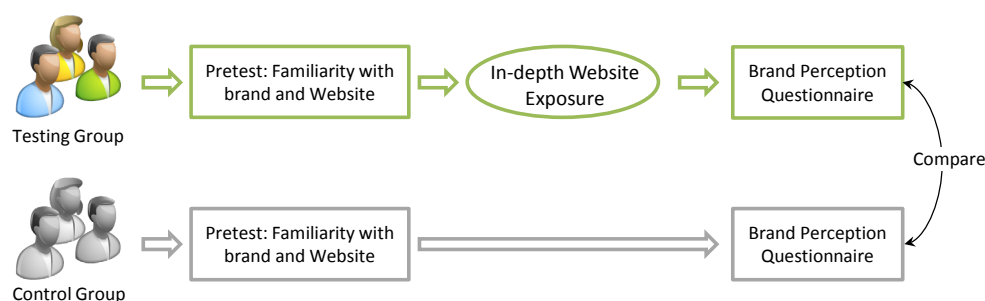


Figure 7. The impact of a website on brand perception can be capture by comparing the brand perception of a testing group (i.e., in-depth exposure to the website) with that of a control group (i.e., no in-depth exposure to the website).

The case study adopted this testing-control group approach in order to evaluate the impact of the AT&T website on users' perceptions of the intended brand identity attributes. In order to get a large sample size, the study was conducted online using a custom designed user testing platform (See the end of this section for more information).

- 2) The ideal situation for investigating a website's impact on its related brand is to recruit people who have never used the website before. This, however, sometimes cannot be achieved because of the low availability of the target users. This is especially the case for well-known brands whose websites have been highly exposed to public audiences. Therefore, it is also necessary to capture users' familiarities with the website under evaluation so that the researchers can exclude the influence of this factor when analyzing the evaluation results. In order to do this, the following item can be added to the pre-test questionnaire (Figure 7):

Item#3. Are you familiar with [name of the brand]'s website (insert the URL of the site)?

Never Used It ----- 1 ----- 2 ----- 3 ----- 4 ----- 5 ----- 6 ----- 7 ----- Very Familiar

Phase Five will show how this factor can be statistically controlled using an ANCOVA test.

- 3) Ideally, all of the users recruited in the evaluation should be the real target users defined in Phase Three. For example, we may ask actual digital TV users to perform the tasks that we designed for digital TV users. However, whether this condition can be met depends upon the availability of the target users and the testing context. For instance, special user groups, such as soldiers or patients, may not be easily accessible. In case researchers want to conduct an evaluation online, in which people voluntarily participate, it is even harder to control the profiles of the users.

According to the nature of the products and services that AT&T offers, however, it is safe to conduct the evaluation online because almost everyone has used TVs, the Internet, home phones, and cellphones.

- 4) The procedure of the evaluation becomes quite straightforward after the decision is made in regard to how prior brand knowledge is going to be measured.

In the situation when prior brand knowledge is measured using a single testing group, users can first be asked to complete a pre-test questionnaire containing Items #1, #2, and #3. Then, they will be given tasks to perform on the website under evaluation. Upon finishing the tasks, they fill in the brand perception questionnaire.

If prior brand knowledge is measured using a control group, then the testing group will first be asked to complete a pre-test questionnaire containing Items #1 and #3 (Figure 7). Then, the testing group will perform tasks and complete the brand perception questionnaire. The control group will only be asked to complete one questionnaire consisting of item #1, #3, and the items from brand perception questionnaire.

All of the users' demographic information should also be collected in case any potential gender or age differences exist. Examples of demographics questions can be found in Appendix C.

- 5) As stated in the beginning of this section, whether the branding effectiveness of a website is high or low depends upon how well the competitors' websites perform within a similar or the same evaluation.

In order to capture the branding effectiveness of a competitor's website, researchers can either conduct a comprehensive evaluation of the competitor's website by going through Phases One to Four (i.e., the same process as following for the main website under evaluation) or, based on the aim of the overall evaluation, the process can be simplified accordingly. For example, researchers might only be interested in seeing how well a competitor's website can communicate the same brand identity attributes that the brand under evaluation is trying to convey. Thus, the same brand perception questionnaire can be applied to the evaluation of the competitor's website, which saves the researchers' efforts in regard to analyzing the intended identity of the competitor brand. However, if the researchers also want to investigate the *unique*

brand identity attributes that the competitor brand is emphasizing, then these unique attributes should also be added into the brand perception questionnaire.

In the case study, the same brand perception questionnaire used to assess AT&T was used to assess the Verizon website in order to see how well the Verizon website could communicate the same brand identity attributes as were communicated by AT&T. The only modification needed was to replace “AT&T” with “Verizon” in the subject of each item. For example, the item, “The products that AT&T provides are innovative” was modified to “The products that Verizon provides are innovative.”

The screenshot shows the AT&T website interface. A red dashed box highlights the main content area, which includes a promotional banner for smartphones and a navigation menu. A red solid box highlights the 'Access Your Account' section, which contains login fields for User ID and Password, and a 'Log In' button. A red solid box highlights the 'Task 2/4' section, which contains a task description and a 'Next' button. A red solid box highlights the 'Progress: [3/46]' bar at the bottom of the task section.

Website Under Analysis: an existing full website is embedded as it is in the window frame.

Ongoing Tasks: tasks are displayed to the participants one by one together with options to fill in with task results.

Navigation Commands: participants can click on the button to go to the next task.

Participants have a clear view of the status of the entire test thanks to the **progress bar**.

Figure 8. Using the OFTEN testing environment to conduct the BREW evaluation online.

As mentioned earlier, the BREW evaluation can be conducted online using a custom designed user testing environment. The name of the environment is OFTEN, which stands for: Online Field Testing ENvironment (Yang, Linder, & Bolchini, 2012). The OFTEN environment was programmed using simple HTML, CSS, and PHP. It is open source so that researchers can easily download and modify it for their own research

purposes.¹ In particular, OFTEN is suitable for the type of user studies that involve pre-/post-test questionnaires and a task performance session. The main interface of OFTEN is shown in Figure 8. It splits a screen into two parts. The top part, which occupies the top 2/3 of the screen, displays the website under evaluation, while the bottom part, which occupies the lower 1/3 of the screen, displays the tasks used for user testing. Only one task is shown to the users at a time. Users can move to the next task by clicking on the “Next” button. In the case study, I made the tasks multiple-choice questions with four answer choices: one right answer, two interference answers, and a “Give Up” option. In case the users could not find the answer, they could simply select the “Give Up” option and click on the “Next” button to skip the task. The environment also provides elements through which to capture the amount of time a user spends on each task (i.e., time-on-task) and templates for building the questionnaires.

Capitalizing on the OFTEN environment and study design, four testing routes were created: 1) users fill in the pre-test questionnaire, perform tasks on the AT&T website, and fill in the post-test brand perception questionnaire (AT&T testing group); 2) users fill in the pre- and post-test questionnaires about AT&T (AT&T control group); 3) users fill in the pre-test questionnaire, perform tasks on the Verizon website, and fill in the post-test brand perception questionnaire (Verizon testing group); 4) users fill in the pre- and post-test questionnaires about Verizon (Verizon control group). The OFTEN environment was able to randomly assign a user to one of the four routes, which guaranteed that each group was made up of different users. Upon finishing the study, the OFTEN environment generated a unique reward code for each user, which will be explained in the next paragraph.

The participants of the case study were recruited using the famous crowdsourcing platform, Amazon Mechanical Turk. Amazon Mechanical Turk accommodates two types of roles: requesters and workers. The requesters can post their tasks on the Amazon Mechanical Turk, while workers can work on the tasks after they are posted and receive monetary rewards upon finishing. In the case study, I created (as a requester) a simple post on the Amazon Mechanical Turk, which only consisted of several lines of instructions and a text box. The post instructed the users to participate in the study by

¹ The source code of the OFTEN environment can be downloaded from:
<http://mypage.iu.edu/~taoyang/research/OFTEN.html>

clicking on a link, which would lead them to the pre-programmed OFTEN environment. It also told users that they would get a reward code after finishing the study that they would need to submit within the text box in order to receive their \$1.50 incentive. This amount is the standard incentive suggested by Amazon Mechanical Turk for 15 minutes of participation.

3.1.5. Phase Five: Result Analysis

How the evaluation results should be analyzed depends upon which types of participants are recruited in the user test and the way in which the users' prior brand knowledge is captured. Researchers may follow the general suggestions below to customize their own result analysis strategies.

- 1) In an ideal condition, researchers are able to recruit participants who have never heard about the brand and its website. Therefore, the branding effectiveness of the website is directly reflected by the ratings that users give in the brand perception questionnaire.
- 2) In a situation in which the participants (or a portion of the participants) already know the brand, but know nothing about its website, the branding effectiveness of the website can be examined through a factorial ANOVA.

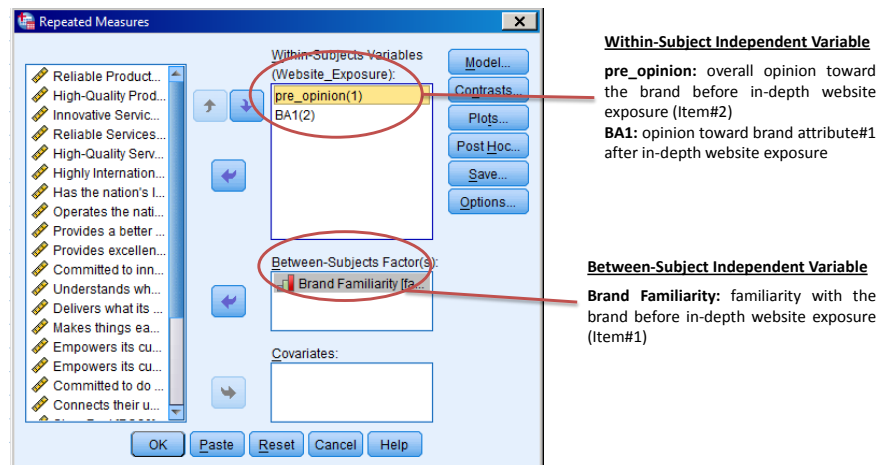


Figure 9. When prior brand knowledge is measured by a single testing group, a mixed-model ANOVA can be conducted to examine the impact of website on brand perception.

The first independent variable in the ANOVA analysis is the users' familiarity with the brand. The users could be split into two or three groups according to their familiarity with the brand (captured by pre-test Item#1). For example, people who select '1' ("Never Heard of It") for Item#1 can be considered *newcomers*, those who select '2' through '4' are *novices*, and those who select '5' through '7' are *experts*.

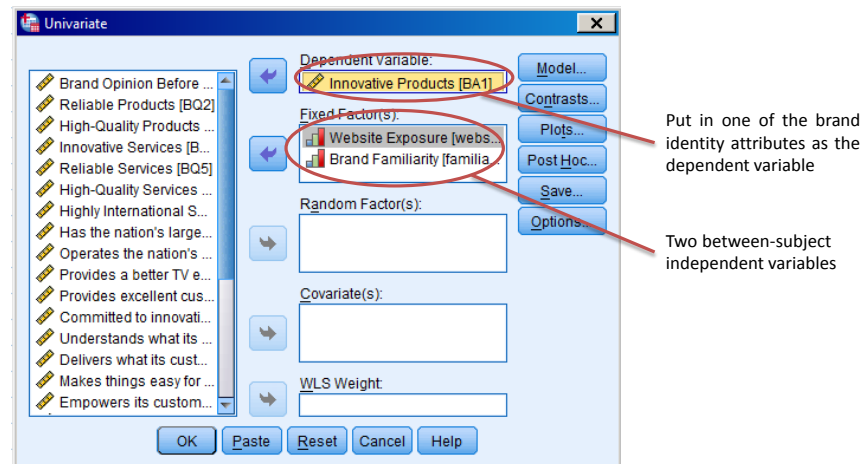


Figure 10. When prior brand knowledge is measured by a control group, a between-subject ANOVA can be conducted to examine the impact of the website on brand perception.

The second independent variable is website exposure, which consists of two conditions: *no in-depth exposure* (i.e., the users did not perform tasks on the website) and *with in-depth exposure* (i.e., the users performed tasks on the website). If prior brand knowledge is measured using a single testing group, a mixed-model ANOVA can be conducted with brand familiarity (Item#1) as a between-subject factor and website exposure as the within-subject factor. The dependent variable of the ANOVA is the users' brand perceptions. The brand perception of the *no in-depth exposure* condition is captured by pre-test Item#2, while the brand perception of the *with in-depth exposure* condition is captured by each one of the items in the post-test brand perception questionnaire. Figure 9 shows how this ANOVA test can be performed using SPSS. Through this test, researchers can get to know, for users with a certain familiarity level with the brand, how much their perceptions of a particular brand identity attribute have been changed by the website in comparison to their overall opinions toward the brand before their in-depth website exposure (Item#2). The limitation of this approach, however, is that the brand perception of the *no in-depth*

exposure condition is only measured by one item (Item#2), which may not accurately reflect the users' opinions toward each brand identity attribute.

If prior brand knowledge is measured using a control group, *website exposure* also becomes a between-subject predictor. In this case, a between-subject factorial ANOVA should be performed (Figure 10). This approach more accurately reveals the change in the perception of each brand identity attribute. However, researchers should check the homogeneity test (Levene's Test) to make sure that the variances of the two groups are equal (Field, 2009).

- 3) If all of the participants or a portion of the participants already know both the brand and its website, then the researchers may need to factor out the influence of the users' familiarity with the website. This can be done through an ANCOVA test by adding website familiarity (pre-test Item#3) as a covariate to the ANOVA tests above (see the "Covariates" box in Figure 9 and Figure 10). The influence of the covariate will be statistically controlled so that the users will be put onto a same website familiarity level before doing all the comparisons. For example, if the evaluation involves testing and control groups, then an ANCOVA can be conducted using *brand familiarity* and *website exposure* as the two between-subject independent variables, *brand perception* as the dependent variable, and *website familiarity* as the covariate. In this manner, researchers can see the pure effect of brand familiarity and website exposure on brand perception without considering the influence of website familiarity.
- 4) The previous three guidelines are used for interpreting the effect of a website on the image of the related brand. In order to decide how strong the effect is, researchers need to compare it with the performances of the competitors' websites. The difficulty of making this comparison, however, is that users of different brands may have different prior knowledge of a particular brand identity attribute. For example, after website exposure, the rating of a brand identity attribute increased from four to six for Brand#1 ($\text{delta} = 2$), whereas for Brand#2, the rating increased from five to seven ($\text{delta} = 2$). Only comparing the rating after website exposure may cause a false conclusion that Brand#2's website is better at communicating the attribute because seven is higher than six. In order to address this issue, in addition to comparing the final brand perception score, researchers should investigate which website causes a

better *change* in brand perception (i.e., the magnitude of the delta). This can be done in two different ways:

If prior brand knowledge is measured using a single testing group, researchers may add users' prior brand opinion (Item#2) as an additional covariate. In other words, the ANOVA test should have *website* (i.e., the website under evaluation and its competitor's website) and *brand familiarity* as the two independent variables and *website familiarity* and *prior brand opinion* as the two covariates. In this way, the researchers can compare the impact of the two websites on brand perception on the basis of equal prior brand opinion.

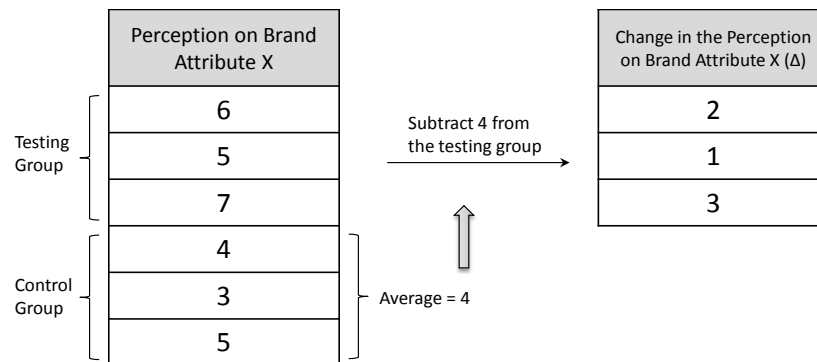


Figure 11. Calculating the change in brand perception.

In contrast, if prior brand knowledge is measured using a control group, then the researchers will first need to calculate the change in brand perception (i.e., the delta). Since the testing and control groups are composed of different participants, the researchers cannot directly subtract the score given by the control group from that of the testing group. Instead, they must first calculate the average score that the control group gives to a particular brand identity attribute, and then subtract this average score from the score that the attribute received from the testing group (Figure 11). The delta score enables researchers to investigate whether the website under evaluation can bring a better change in brand perception than its competitors' websites.

Following the guidelines above, the results of the case study are analyzed in detail in the next section.

3.2. Results of the Case Study

A total of 216 complete responses (i.e., finished the entire study and submitted the reward code) were received from Amazon Mechanical Turk. Then, the data was polished by removing the responses that showed apparent evidence of gaming the testing system as well as the extreme outliers. Eventually, 202 cases were retained: 50 belonging to the AT&T testing group, 59 to the AT&T control group, 42 to the Verizon testing group, and 51 to the Verizon control group. Although there were more participants in the Verizon testing group who were not able to accomplish the study, a Chi-Square test did not suggest significant difference in terms of the amount of valid responses that each group contained. Among the 202 participants, 85 were female and 117 were male. Most of their ages were within the 18 to 40 range ($N = 182$), while the rest were above 40 years old ($N = 20$). Most of the respondents were pursuing or had received an undergraduate degree ($N = 156$), while the rest either had not attended college ($N = 27$) or went to graduate school ($N = 19$). I re-grouped the participants according to their age and education in order to make sure that each group contained an approximately equal number of participants.

3.2.1. The Impact of the AT&T Website Itself

Overall, the average score that the AT&T control group gave to each brand identity attribute was higher than four (out of seven), which is over the neutral point. This implied that, before all of the users used the AT&T website in-depth, they had, in general, a relatively positive attitude toward AT&T. The three brand identity attributes that received the lowest average scores were: [AT&T] operates the nation's largest Wi-Fi network ($M = 4.29$, $SE = .15$), provides a better TV experience ($M = 4.31$, $SE = .15$), and [AT&T users are] saving [money] ($M = 4.31$, $SE = .19$).

The AT&T testing group also gave positive average scores to AT&T ($M > 4$) after all users used the AT&T website in-depth. The three brand identity attributes that received the lowest average scores were: [AT&T is] caring [for its customers] ($M = 4.06$, $SE = .26$), [AT&T is] righteous ($M = 4.22$, $SE = .23$), and [AT&T users are] saving [money] ($M = 4.24$, $SE = .22$).

The data showed that most of the participants were highly familiar with AT&T. Across the control and testing groups, 100 participants gave a rating of at least five (out of seven)

regarding their familiarity with AT&T (high familiarity), whereas only nine participants gave a rating equal to or lower than four (low familiarity). Due to the unbalanced sample size, instead of conducting an omnibus ANOVA, t-tests were performed for the high and low brand familiarity groups separately. The t-tests had *website exposure* (two levels: no in-depth exposure and with in-depth exposure) as the independent variable, the *perception on one of the brand identity attributes* as the dependent variable, and *website familiarity* as the covariate.

The t-tests detected two significant differences in brand perception between the testing and control groups (Figure 12). For people with low familiarity with AT&T, the testing group gave a significantly higher rating to the brand identity attribute *saving* ($M = 5.6$, $SE = .40$) than the control group ($M = 2.76$, $SE = .46$), which did not have in-depth exposure to the AT&T website ($F(1, 6) = 18.01$, $p < .01$). In fact, the control group's rating was already leaning toward the "wasting" side. For people with high familiarity with AT&T, in contrast, the testing group's rating of *valuable* ($M = 4.48$, $SE = .25$) was significantly lower than the control group's rating ($M = 5.17$, $SE = .23$), indicating that the AT&T website might have a negative impact on the communication of this attribute ($F(1, 97) = 4.12$, $p < .05$).

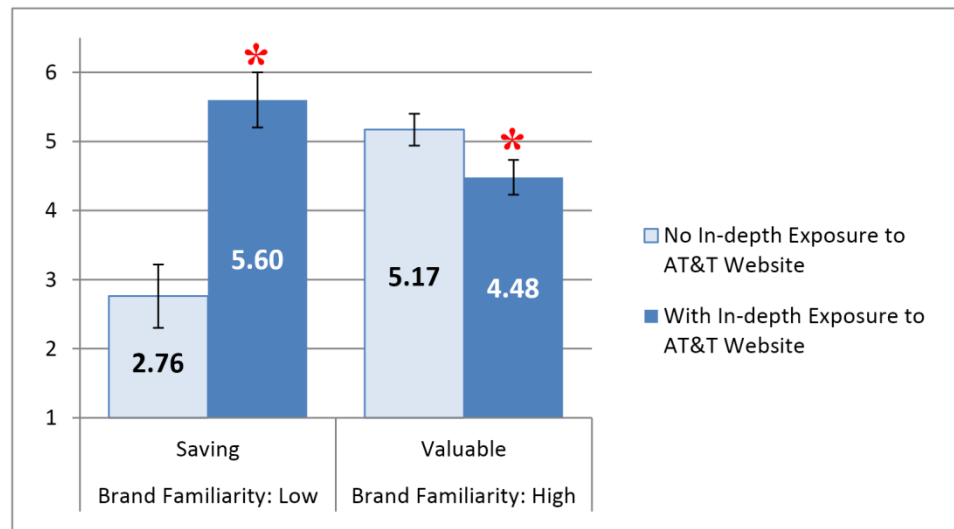


Figure 12. In-depth exposure to the AT&T website brought two significant changes in brand perception (*significantly different from the control group, $p < .05$).

For the high brand familiarity group, significant gender differences were found by checking the interaction of *website exposure* * *gender* through factorial ANOVAs. The most salient findings are illustrated in Figure 13. In spite of the statistical significance, we can see that females' and males' brand perceptions changed in opposite directions after their in-depth website exposure. In general, females' ratings of the brand identity attributes in Figure 13 decreased, while males' ratings increased. This finding indicated that the AT&T website might contain elements that only females did not like, which negatively affected their brand perceptions. In particular, females' perceptions significantly decreased in regard to five attributes: reliable products, delivers what its customers want, caring, righteous, and reliable. The perceptions of *caring* and *righteous* were already on the negative side ($M < 4$) after the in-depth website exposure. On the contrary, males' perceptions significantly increased in four attributes after their in-depth website exposure: innovative products, high-quality products, better TV experience, and delivers what its customers want. The perceptions of *better TV experience* and *delivers what its customers want* actually increased from negative to positive (from lower than four to higher than four).

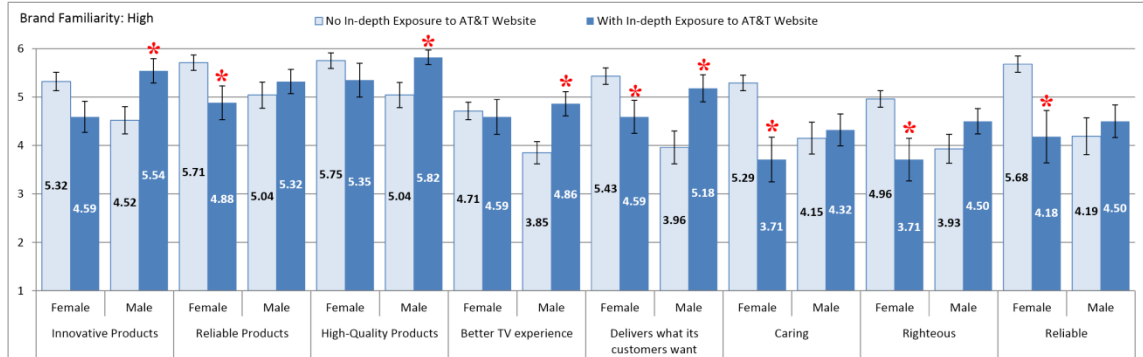


Figure 13. The AT&T website may have a different impact on females and males in regard to brand perception (*significantly different from the control group, $p < .025$).

In addition to gender, I checked the effects of other demographics traits (i.e., age and education), but no significant differences were found.

3.2.2. The Comparative Branding Effectiveness of the AT&T Website

Similar to the AT&T portion of the study, the Verizon control group held positive attitudes toward most of the brand identity attributes. The only exception was the attribute *saving*, which received a negative rating ($M = 3.94$, $SE = .24$). The Verizon testing group, in

contrast, gave positive ratings to all of the brand identity attributes, including the attribute *saving* ($M = 4.50$, $SE = .23$). T-tests comparing the ratings from the Verizon control and testing groups (with website familiarity statistically controlled) revealed that, for people with low familiarity with Verizon, their perceptions of two brand identity attributes might improve significantly after in-depth exposure to the Verizon website (Figure 14): *empowers its customers to use technology safely* ($F(1, 2) = 94.08$, $p < .05$) and *empowers its customers to use technology securely* ($F(1, 2) = 18.75$, $p < .05$). For people with high familiarity with Verizon, however, their perceptions of *reliable services* might significantly worsen after in-depth website usage ($F(1, 85) = 9.97$, $p < .01$). Significant interaction of *website exposure * gender* was found on the perception of *provides excellent customer care* ($F(1, 83) = 4.04$, $p < .05$). Females gave it a significantly lower rating ($M = 4.33$, $SE = .39$) after their in-depth website exposure when compared to the control group ($M = 5.44$, $SE = .28$). Males' ratings, however, were approximately the same between the control ($M = 4.70$, $SE = .33$) and testing group ($M = 4.84$, $SE = .29$).

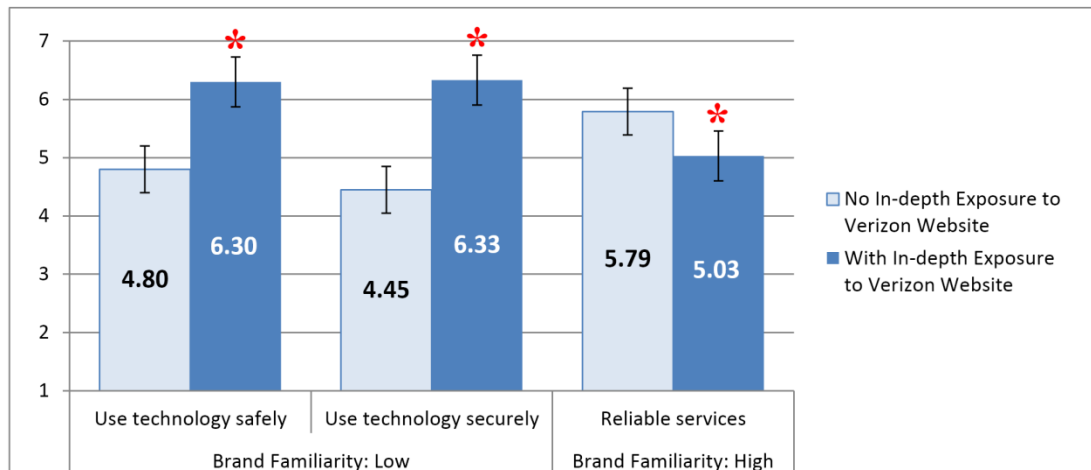


Figure 14. In-depth exposure to the Verizon website brought three significant changes in brand perception (*significantly different from the control group, $p < .05$).

In the situation when the users had high brand familiarity, significant differences were found between the AT&T and Verizon control groups in regard to the perception of four attributes: *reliable services*, *high-quality services*, *has the nation's largest 4G network*, and *makes things easy for its customers*. All four attributes were significantly better perceived by the Verizon control group than the AT&T control group (Figure 15).

However, after in-depth website exposure, all of these significant differences disappeared. In general, compared to the corresponding control group, the perceptions of all four attributes decreased substantially in the Verizon testing group, whereas the perceptions of these attributes stayed almost the same in the AT&T testing group. As a consequence, the ratings of these attributes became much closer between Verizon and AT&T testing groups. No significant differences were found in the other attributes between the AT&T and Verizon testing groups (Appendix D).

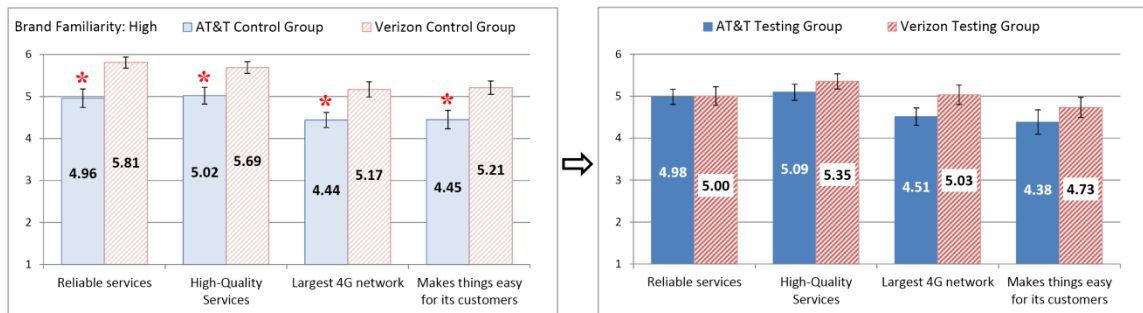


Figure 15. The significant differences in brand perceptions between AT&T and Verizon diminished after in-depth website exposure (*significantly different from Verizon, $p < .05$).

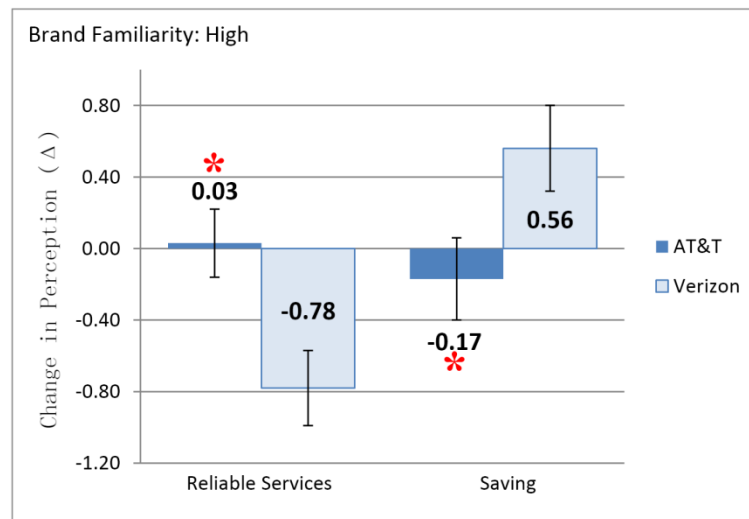


Figure 16. The impact of the AT&T website on the perceptions of two brand identity attributes was different from that of the Verizon website (*significantly different from Verizon, $p < .05$).

Then, following the guidelines in the previous section, the changes in brand perception by website exposure were calculated. For people with low and high brand familiarities separately, t-tests were conducted with the *change in the perception of a brand identity*

attribute as the dependent variable, *website* (two levels: AT&T and Verizon website) as the independent variable, and *website familiarity* as the covariate. The results suggested two significant differences in the high brand familiarity group. First, the impact of the AT&T website on the perception of the attribute *reliable services* was significantly higher than that of the Verizon website ($F(1, 82) = 8.10, p < .01$). After in-depth website exposure, AT&T users' perceptions of *reliable services* on average increased by .03 (SE = .19), whereas that of the Verizon users decreased by .78 (SE = .21) (Figure 16). Second, the impact of the AT&T website on the perception of the attribute "saving" was significantly lower than that of the Verizon website ($F(1, 82) = 4.73, p < .05$). After in-depth website exposure, AT&T users' perceptions on *saving* on average decreased by .17 (SE = .23), whereas that of the Verizon users increased by .56 (SE = .24). No interaction effect was detected based on the users' demographics traits.

3.3. From Evaluation Results to Strategies for Improvement

To summarize the positive aspects from the evaluation results, the AT&T website was able to positively communicate most of the intended brand identity attributes. In particular, for people who were not familiar with AT&T, the website changed their perceptions of the images of AT&T users from *wasting* to *saving*. For males who were highly familiar with AT&T, the website improved their perceptions on four attributes: innovative products, high-quality products, better TV experience, and delivers what its customers want. In addition, it performed significantly better than the Verizon website in conveying the attribute *reliable services*.

The defects of the AT&T website were all revealed by the high brand familiarity group. First, the perception of the attribute *valuable* was significantly worsened after in-depth website exposure. Second, for females, their perceptions on five attributes significantly worsened after in-depth website exposure: *reliable products*, *delivers what its customers want*, *caring*, *righteous*, and *reliable*. Third, the AT&T website's impact on the attribute *saving* was significantly worse than that of the Verizon website. Finally, the differences between AT&T and Verizon in regard to the perceptions of four brand identity attributes diminished after in-depth website exposure (Figure 15). The disappearance of the differences, however, was not because the AT&T website improved the perceptions of these attributes, but because the Verizon website worsened them. Therefore, if time and

budget allowed, AT&T may also need to enhance the communications of these four attributes, but it is not imperative to do so.

The following sections show how improvement strategies can be made to the first three issues. All of the analyses and discussions made below are focused on the high brand familiarity group.

3.3.1. Strategies for Enhancing the Communication of the Attribute “Valuable”

In the original term, AT&T stated that “[we want to] ensure we are as valuable to our customers as they are to us.” In order to interpret why the website had a negative impact on this attribute (Figure 12, the perception of the attribute decreased 13.35% after in-depth exposure), we first need to check the branding-design map (Figure 5) to see where this attribute was conveyed in the selected website sections. According to the map, the attribute was communicated in 17 webpages. Eight of these pages explicitly communicated the attribute. For example, in the “U-verse TV” page, it says “it [the U-verse TV] comes at *a great value*.” Twelve of the pages implicitly communicated the attribute. For example, in the “AT&T Wi-Fi” page, it says “AT&T Wi-Fi allows users to access the Internet while on the go at thousands of hotspots nationwide,” which *adds value* to regular Internet services.

In the user test, in order to accomplish the four task scenarios (Appendix B), the users at least needed to browse seven of the 17 pages (6 explicit and 1 implicit), which should have given them a thorough exposure to the attribute. However, they were not able to strongly perceive these attributes for two possible reasons. First, the two tasks relating to the High Speed Internet and Home Phone sections of the AT&T website yielded quite low success rates (42.22% for the former and for 37.78% for the latter). These two sections, however, embraced four of the above seven pages where the attribute *valuable* was conveyed. If the users were not able to get to these pages, there was no way for them to perceive the attribute. Second, the attribute *valuable* was not emphasized in any of the relevant pages. It was either embedded in the regular text or in a place below the fold (i.e., users needed to scroll down to see it), which users can easily overlook (Figure 17).

Based on the analyses above, two recommendations can be made.

- 1) Improve the usability of the High Speed Internet and Home Phone sections. Since the study was conducted online, we do not know exactly why most of the users could not accomplish the two tasks related to the two sections. Therefore, researchers may conduct a small scale lab-based usability test to observe what problem users may have in finding information from the two sections.
- 2) Make the attribute more easily perceivable. On the one hand, the attribute can be moved above the fold on the webpages so that users can see it without scrolling down. On the other hand, designer may use larger font or create banners or slogans to draw users' attention.

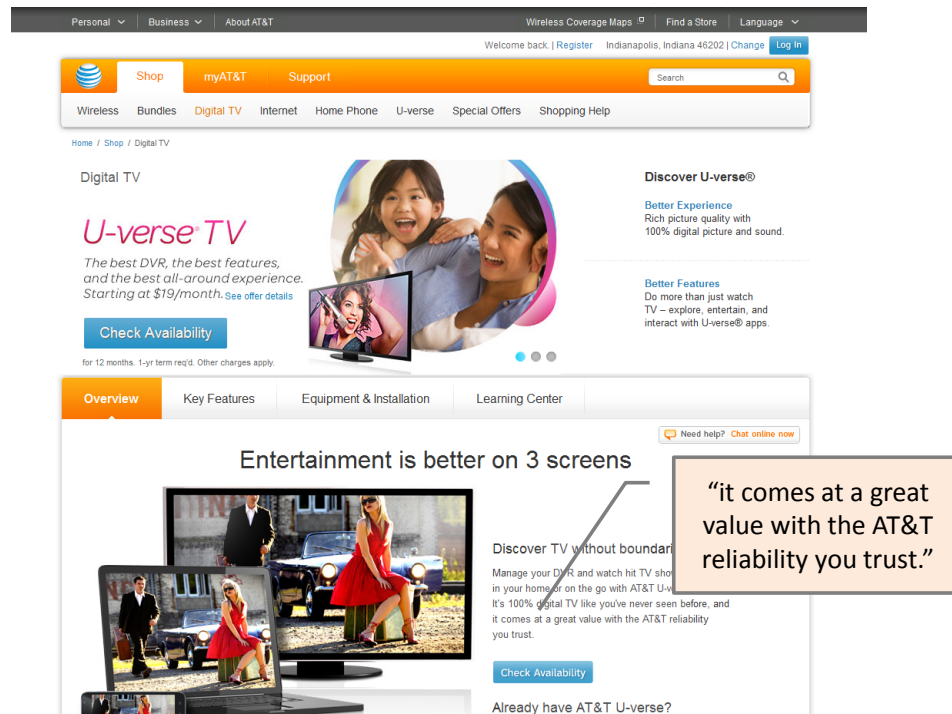


Figure 17. With very small font and a not so important location, the attribute “valuable” is not emphasized enough on this page.

3.3.2. Possible Improvement Strategies for Female Users' Brand Perceptions

Only females gave significantly lower ratings to five of the brand identity attributes after thoroughly using the AT&T website (Figure 13). The branding-design map showed that two of these attributes (i.e., *reliable products* and *righteous*) were not pervasively communicated in the website. They each only appeared twice in the selected website

sections. The low presence, however, does not seem to be the only reason why only females gave much lower ratings to these five attributes because three other attributes also worsened even though they had quite a high presence on the website. Moreover, usability was also unlikely to be the reason because females' ratings on these attributes were not significantly correlated with their task success rates and time-on-task.

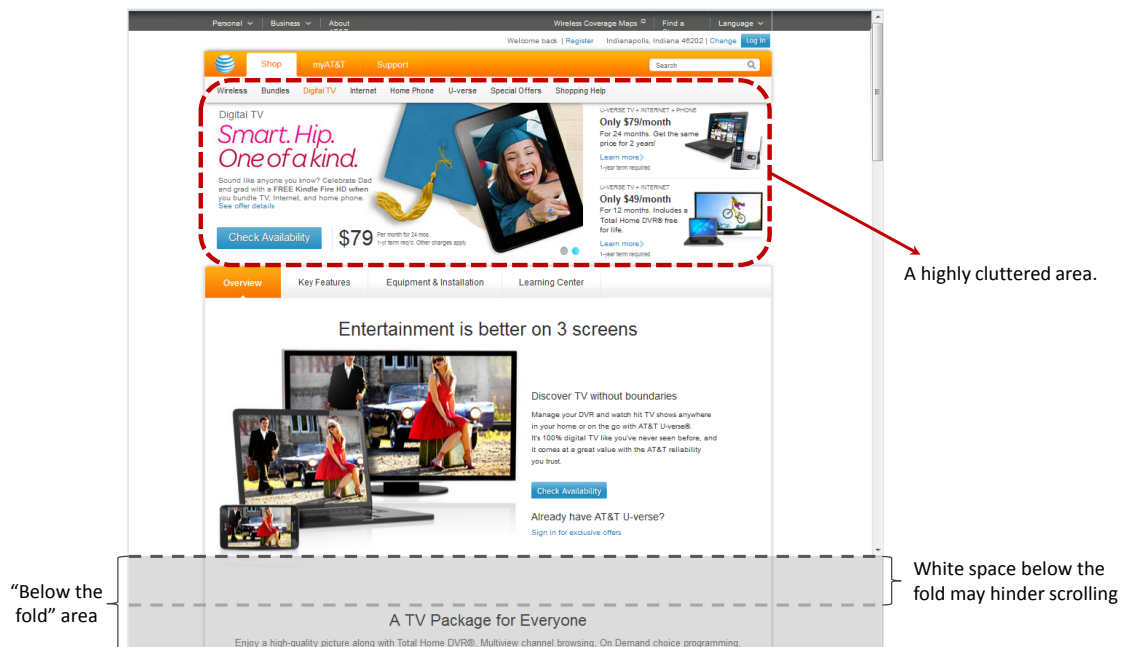


Figure 18. Visual guidance problems of the AT&T website may negatively affect females' perceptions on several brand identity attributes.

In fact, this type of gender difference in online brand perceptions also showed up in another study that I conducted (Chapter 5). That study revealed that females' online brand perceptions were sensitive to the *visual guidance* design of a webpage, which has to do with the highlighting and positioning of the webpage elements. Poor visual guidance (i.e., cluttered layout, placing an important element in a trivial position, or scrolling stoppers) could cause females to give poor ratings to a number of brand identity attributes of a US state university, but males were not affected. Accordingly, I inspected the webpages related to the five worsened AT&T attributes in order to see if these worsened ratings about AT&T were caused by the same problem. Coincidentally, many of these pages were indeed cluttered, especially the upper part of the page where the designers cluttered the page with texts in different font sizes and colors as well as a number of images (Figure 18). Moreover, some of the long pages used wide white

spaces to divide sections horizontally, which may have prevented the users from scrolling down (scrolling stoppers, Nielsen, 2006). In contrast, the pages of the Verizon website had much cleaner and pleasing visual guidance designs (Figure 19). Therefore, the following two recommendations can be made.

- 1) Address the visual guidance issues of the pages related to the communication of the five brand identity attributes. In particular, the layout-clutter issues and scrolling stoppers should be removed. Since these issues only affected females, researchers may need to test the improved pages with female users to confirm the effectiveness of the redesign.
- 2) Strengthen the communication of the two less pervasively conveyed brand identity attributes (i.e., *reliable products* and *righteous*). Embed them in more webpages related to typical user task scenarios.

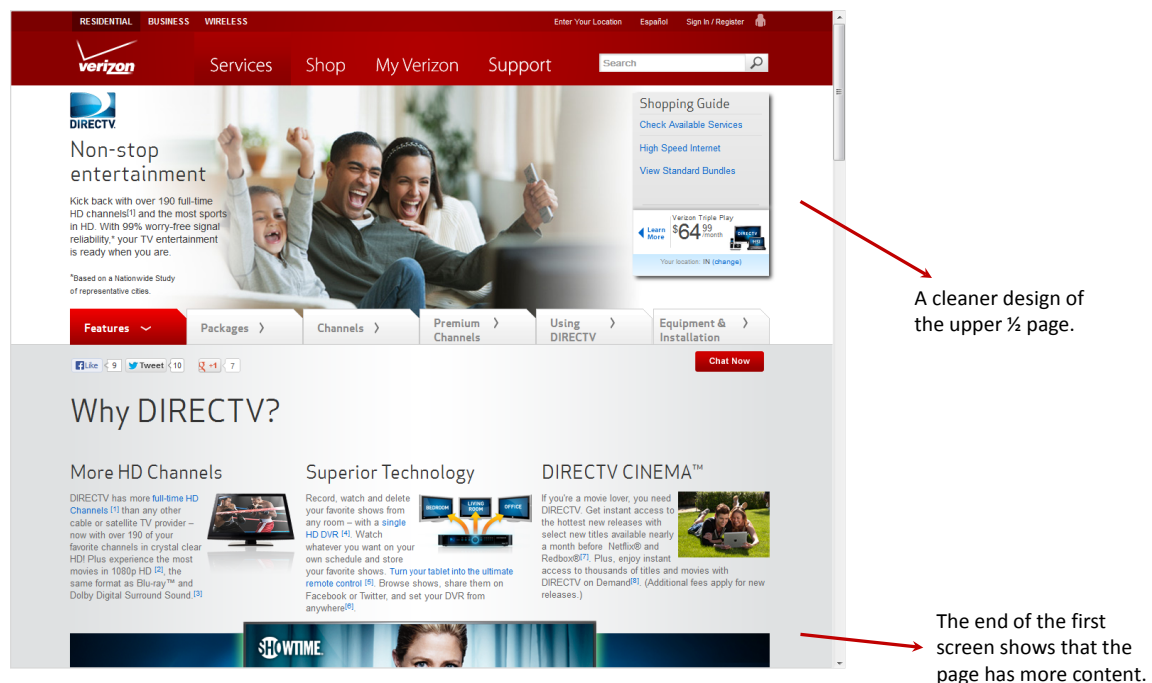


Figure 19. A better visual guidance design by Verizon.

3.3.3. Strategies for Enhancing the Communication of the Attribute “Saving”

The Verizon website had a significantly better impact on the communication of the attribute *saving* than the AT&T website (Figure 16). Although the AT&T website tried to

convey, on almost every page, that users do not need to spend much money to use a service or product, the average rating that the users gave to this attribute after using the website was only 4.13 (SE = .33), which was quite close to the neutral point. This rating was not significantly lower than the rating given by the AT&T control group (M = 4.38, SE = .20), but, at least, it meant that the AT&T website was not helpful in enhancing the communication of the attribute *saving*. In contrast, the average rating (M = 3.90, SE = .25) suggested that the Verizon control group did not think that the Verizon users were saving money. However, after in-depth website exposure, users' opinions became much more positive (M = 4.50, SE = .22).

By comparing the related webpages of the AT&T and Verizon website, I found that the two websites, in fact, used different strategies to communicate the attribute. Most of the time, the message conveyed by AT&T was: *the price is low*. Examples are: "TV as low as \$19 a month when you bundle with Internet" or "Only \$49/month." In contrast, Verizon directly told users: *you save money*. Examples are: "Get \$5 off the monthly price for 2 years when you order a Freedom Plan online" or "You can receive a \$4.99 monthly discount off the international plan." Based on the users' ratings, it is apparent that the latter strategy is more effective. Therefore, the recommended improvement strategy for AT&T is to communicate the attribute *saving* in a more straightforward manner. They may use absolute discount values (like Verizon did), percentages, or comparisons (with the former or competitor's price) to directly demonstrate that users can save money, rather than only telling users that the price is not high.

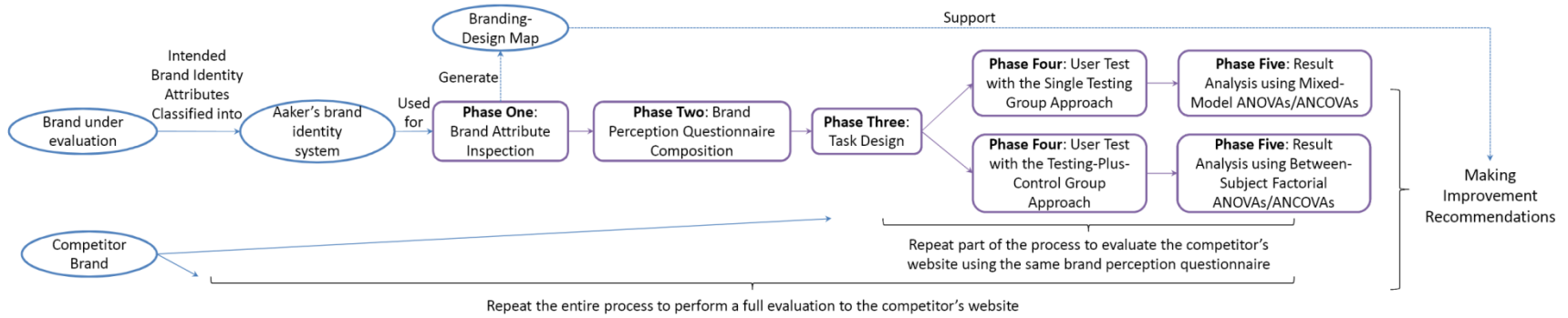


Figure 20. A synopsis of the BREW evaluation approach.

3.4. Discussions

3.4.1. Synopsis of the BREW Evaluation Approach

Figure 20 provides a procedural synopsis of the BREW Evaluation Approach. EBranding researchers can use this procedural chart to get a bird's-eye view of the approach and follow the steps to plan and execute their own evaluation projects. When needed, they can read the detailed guidelines of a specific step in section 3.1. In summary, with the intended brand identity attributes classified into Aaker's branding system by the marketing team, eBranding researchers can perform a *brand attribute inspection* to map the brand identity attributes to the architecture of the website. Next, they can transform the classified brand identity attributes into a brand perception questionnaire. Third, user tasks should be created based on typical user profiles. In the fourth step, researchers need to decide whether to use a single testing group or a testing-plus-control group experiment design. The former requires a smaller sample size, but may not accurately reveal how much each attribute can be affected by in-depth website exposure. The latter is more rigid and captures the change in each attribute, but requires a larger sample size. Researchers also need to decide how the competitor's website should be evaluated. They may either simply go back to phase three to create equivalent user tasks for the competitor's website and continue to use the same brand perception questionnaire, or go through phases one to three again to perform a full evaluation of the competitor's website. The former approach allows researchers to examine how the competitor's website can communicate the same brand identity attributes as the website under evaluation, whereas the latter enables researchers to add in the competitor's specific brand identity attributes to the questionnaire.

If researchers adopt the single-testing group approach, then they can perform mixed-model ANOVAs/ANCOVAs to examine how each brand identity attribute is affected by in-depth website exposure when compared to an overall prior brand opinion. If they adopt the testing-plus-control group approach, then between-subject factorial ANOVAs/ANCOVAs can be performed to see the change in each brand identity attribute compared to the users' prior opinions toward each specific attribute. Finally, recommendations for improvement should be made based on the evaluation results. Researchers can trace a poorly communicated brand identity attribute back to the branding-design map in order to determine which webpages were used to convey the attribute and why the communication was not satisfactory.

3.4.2. Online or Lab-Based Evaluation

Capitalizing on the OFTEN field testing environment and combining it with Amazon Mechanical Turk, eBranding researchers are empowered to conduct the BREW evaluation online. Based on my experiences, for a 20- minute experiment, it takes researchers about an hour to finish testing over 100 participants. Moreover, in a situation in which the target users' profiles cover a broad range of the population (e.g., with AT&T), Amazon Mechanical Turk allows researchers to get access to participants with much more diversified backgrounds across the global rather than conveniently sampling local users (Nelson & Stavrou, 2011). This advantage, however, may become a limitation if researchers want to strictly control for the participants' backgrounds. For example, a pharmaceutical company may only want pharmacists to participate in the evaluation. It is true that researchers can set up screening criteria to prevent unqualified users from joining the study, but cheating is not easily preventable. Also, people may try to game the testing system to get the incentive without actually performing the tasks (see Section 3.4.4 for the strategies that I used to prevent this action from happening). Moreover, researchers cannot easily get in touch with the participants to get a sense of why they give a low/high rating to a particular brand identity attribute.

In contrast, researchers may want to conduct the BREW evaluation in a usability lab if they need a better control over the users' profiles and the evaluation process, and they would like to observe or interview the users in order to collect more qualitative feedback (Sharp, Rogers, & Preece, 2006). Qualitative feedback can be used to better interpret quantitative findings and uncover users' reasoning regarding how they relate their perceptions of a website to the particular aspects of brand image. The disadvantages of lab-based evaluation, however, are the high costs of inviting users to come to the usability lab and the amount of time and effort researchers spend to facilitate the user tests. Moreover, researchers primarily draw samples from the local population, which may affect the generalizability of the evaluation results.

A possible way to circumvent the limitations of both approaches is to first conduct a large scale evaluation online in order to extract some general quantitative trends. Then, for the findings which needs further interpretation, a smaller scale, lab-based follow-up evaluation can be conducted for more in-depth investigations.

3.4.3. Validity and Limitations of the BREW Approach

First, the BREW evaluation approach is built upon the well-known Aaker's brand identity planning model. Capitalizing on the model, researchers can have a clear understanding of the nature of a particular brand identity attribute: whether the attribute is describing the brand's product, the organization who owns the brand, the personality of the brand, and/or the image of the users (Aaker, 1996). With this understanding, researchers can come up with a more accurate strategy by which to measure a particular brand identity attribute in the brand perception questionnaire, instead of simply using a same scale to gauge all of the attributes.

Second, previous versions of the BREW approach (Bolchini, Yang, & Garzotto, 2009; Yang & Bolchini, 2010) focused more on the preparation of the user testing instruments than on the user study design and result analysis parts, which had not yet been richly developed. The current version substantially advanced these two parts by providing detailed guidelines and insights for eBranding researchers to use when reflecting on the objectives of their own projects and deciding which experiment design and data analysis techniques to adopt. Moreover, the current BREW approach integrated the common practice in marketing (Ambler, 2003; De Chernatony, 2006; De Chernatony & McDonald, 2003) to assess the performance of the website under evaluation by comparing it with that of its main competitor's website. This approach could offer more practical insights (e.g., understand whether the brand has fallen behind its main competitor on certain aspects) than looking at the website under evaluation alone.

Third, the case study demonstrated the effectiveness of the BREW approach. On the one hand, the evaluation results identified the brand identity attributes that might be vulnerable to the current AT&T website. Then, using the branding-design map, which was the result of the *brand attribute inspection* (Phase One), the webpages that conveyed those attributes were retrieved in order to examine the possible reasons why these pages yielded negative impacts. Accordingly, practical improvement recommendations were made for addressing the identified issues. On the other hand, through comparisons with the Verizon website, the brand identity attributes that had *comparative disadvantages* for AT&T were elicited. Although these attributes were not negatively affected by the current AT&T website, the online communication strategies should be improved because the main competitor was doing better.

Finally, in the case study, the improvement recommendations were made based on in-depth analyses of the features of particular webpages. However, there was no way to ensure that the decisions and recommendations made were proper except if the company could implement the recommendations and test the website again. This limitation is caused by the scarcity of empirical studies investigating the relation between website designs and branding effectiveness. Only if eBranding researchers develop richer knowledge in regard to understanding which type of design can better support the communication of a particular brand identity attribute, can we have a better sense of what to do when detecting a poorly communicated attribute. Chapter 5 provides an icebreaker to this important matter.

Moreover, although the BREW approach showed good validity in predicting and analyzing the branding effectiveness of two telecommunication companies' websites, its generalizability needs to be further examined by applying it to the evaluation of branded websites in other domains.

3.4.4. Validity and Limitations of the Case Study

The case study applied the BREW approach to the evaluation of the AT&T website, which not only demonstrated the feasibility of the approach, but also offered rich examples by which to show how each evaluation phase can be executed. Since the case study was conducted online using Amazon Mechanical Turk, a number of measures were taken to prevent system gaming. First, an incentive of \$1.50 was offered for the less than 15- minute participation, which is slightly higher than the suggested reward. Second, a warning message was posted on the starting page indicating that users who gamed the system would not be eligible for the reward. Third, only the users who have an approval rate (i.e., the percentage of successfully participated in studies) of over 95% were allowed to participate in the study. Fourth, task success and time-on-task were tracked using the OFTEN system. Users who failed too many tasks or finished the entire study in a very short time were rejected. Fifth, responses with obvious signs of system gaming (e.g., selected the same answer for all of the questions) were also excluded (Kittur, Chi, & Suh, 2008).

The case study is limited in two aspects. First, the brand identity attributes of AT&T were elicited from the company's published materials (i.e., the company's profile, value

statement, and TV commercials) and then classified into Aaker's system by the author. Thus, AT&T's brand identity system used in the case study may not be completely accurate. In real world scenarios, the brand identity system should be composed by the company's marketing team, not a third-party research group. However, the primary objective of this case study was to illustrate the elements of the evaluation approach rather than providing real improvement recommendations to the company. Second, the tasks used in the case study were all fact-finding tasks (i.e., asking users to find a piece of information from the website, Kellar, Watters, & Shepherd, 2006). Transaction type tasks, such as paying a phone bill, were not included because of the limited permissions of these features. Because of this shortcoming, users may not be able to have a comprehensive experience with the website. Real world projects should allow users to perform all types of tasks related to their profiles.

Chapter 4. DEEP: Design-oriented Evaluation of Perceived Website Usability

The lack of good evaluation-design alignment is the main limitation of the existing perceived usability evaluation instruments (Chapter 2, section 2.4). In order to compensate for this shortcoming, this chapter first introduces a set of analytical website design dimensions extracted from current web engineering theories and empirical studies. Then, using these design dimensions as the blue print, a series of sub-studies were conducted in order to collect, rephrase, and validate the appropriate perceived measurement items under each dimension. These measurement items were designed in order to communicate the meaning of a design dimension in user-understandable language, so that users could directly evaluate the quality of these design dimension based on their usage experiences. Finally, the outcome of these sub-studies is a standardized psychometric instrument that can reliably capture five design-oriented perceived website usability dimensions: content, structure and information architecture, navigation, layout consistency, and visual guidance (Yang, Linder, & Bolchini, 2012). The instrument was named DEEP: DDesign-oriented Evaluation of Perceived website usability.

4.1. Website Design Dimensions

A design dimension is an orthogonal concern of design decisions of a complex web and hypermedia system, and has been deeply explored and used in hypermedia design methods, web design and requirements engineering models and practice over the last decade (Bolchini & Mylopoulos, 2003; Bolchini & Paolini, 2004; Bolchini & Paolini, 2006; Bolchini, Randazzo, & Paolini, 2003; Conallen, 2003; Garzotto & Paolini, 1993; Lange, 1996; Rossi, Schwabe, & Guimaraes, 2001; see Figure 21). The following design dimensions have been identified as relevant for the initial construction of DEEP: content, structure and information architecture, navigation, labeling, graphics and technological performance. It is important to note that design dimensions can be characterized in different ways and styles and at different levels of granularity according to the purpose of the analysis. In general, the value of the proposed dimensions as defined here is the fundamental property of structuring the design activity around design concerns and design expertise.

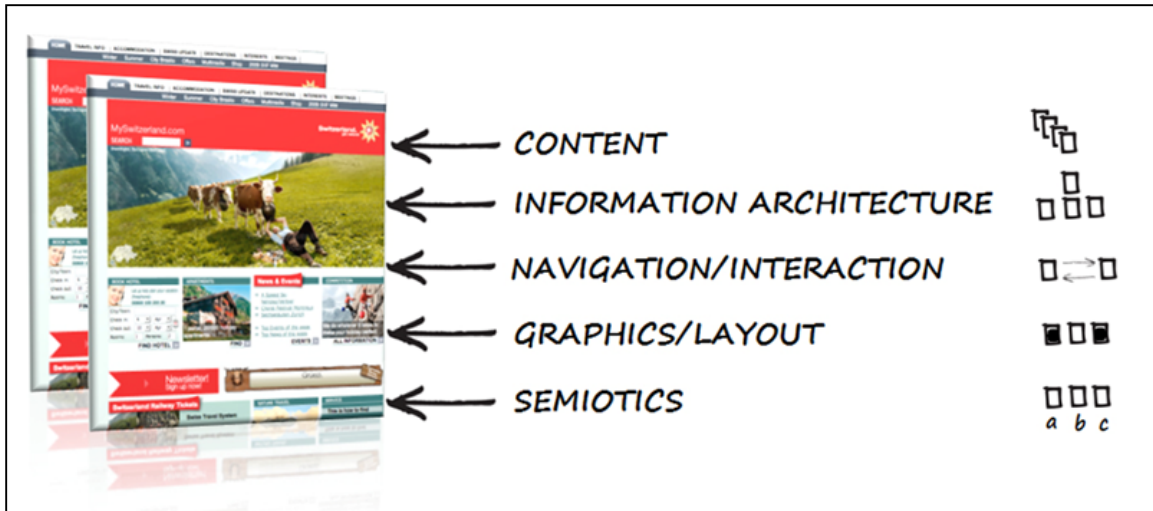


Figure 21. Hypermedia design dimensions play a pivotal role in structuring the requirements, design and evaluation activity (Bolchini & Chatterji, 2009; Bolchini & Paolini, 2004).

Each of the considered design dimensions is discussed in the remainder of this section.

4.1.1. Content

In information-intensive applications, content is the most valuable asset of user experience and represents a fundamental dimension of concern for the success of product design. In this context, content refers to the set of ideas, values, messages, and information that the site communicates to its target audience. Content here is intended in multimedia sense, and includes the asset of textual information, pictures, and videos that are aimed at conveying the key informative messages of the application.

Besides being crucial to requirements analysis and design activity, the evaluation of the usability of content has been explored in the literature to a certain extent. Expected quality attributes of content have been identified for the purpose of usability evaluation. These attributes include relevance, accuracy, currency, coverage, utility, and understandability (Baierova, Tate, & Hope, 2003; Bolchini & Garzotto, 2008; Bolchini & Paolini, 2004; Choi, Lee, & Kim, 2006; Tan & Wei, 2006).

4.1.2. Structure and Information Architecture

The structure and information architecture dimension concerns the way in which a website's content is organized. On one hand, effective and understandable structures organize the content meaningfully according to the key requirements of the application,

user's tasks, and key messages to convey. Poor structures, on the other hand, scatter relevant information for a task throughout the website or 'hide' information inside complicated and non-intuitive classification schemes (Baierova, Tate, & Hope, 2003; Bolchini & Garzotto, 2008; Bolchini & Paolini, 2004; Choi, Lee, & Kim, 2006; Tan & Wei, 2006). This dimension, therefore, concerns those design decisions focused on dividing content into reasonable categories or sections, which could be easily learned and memorized by users, and that can support users in efficiently locating and accessing the needed content.

4.1.3. Navigation

The navigation dimension refers to the design decisions concerned with providing users with appropriate paths and interaction mechanisms to effectively use and move within the information architecture (Baierova, Tate, & Hope, 2003; Bolchini & Garzotto, 2008; Bolchini & Paolini, 2004). On top of a hierarchical information architecture, for example, designers may design flexible navigation patterns to skip hierarchical levels, provide users with accelerators to easily move up and down the hierarchy levels, or filter elements of the collection (Choi, Lee, & Kim, 2006; Ruddle, 2009; Tan & Wei 2006). Good navigation can help users quickly find the information they need and move from one section to other, related sections. By contrast, users may easily get lost or be confused by poor navigation strategies (Park, 2007; Petrie et al., 2009).

4.1.4. Labeling

Labeling is the most *semiotic* of the design dimensions and indicates decisions concerned with the nature of the *interface signs* to use to convey the presence and characteristics of the content, information architecture, and navigation mechanisms. In other words, labeling refers to the language of the interface (Andersen, 2001; Bolchini & Chatterji, 2009; Leite, 2002; Scolari, 2009). By comparing users seeking information to "animals hunting their prey" (p.1), Spool, Perfetti, and Brittan (2004) suggested providing users with the *scent* when they arrive at a website, which could guide them step-by-step to the information they need. According to the authors, labels of the links or buttons on a website are the most effective channels for delivering the scent. Good labels tell users where to go next or what operations pushing a button will perform. However, users may get frustrated if the wording of the label is confusing or is not consistent with the content it represents.

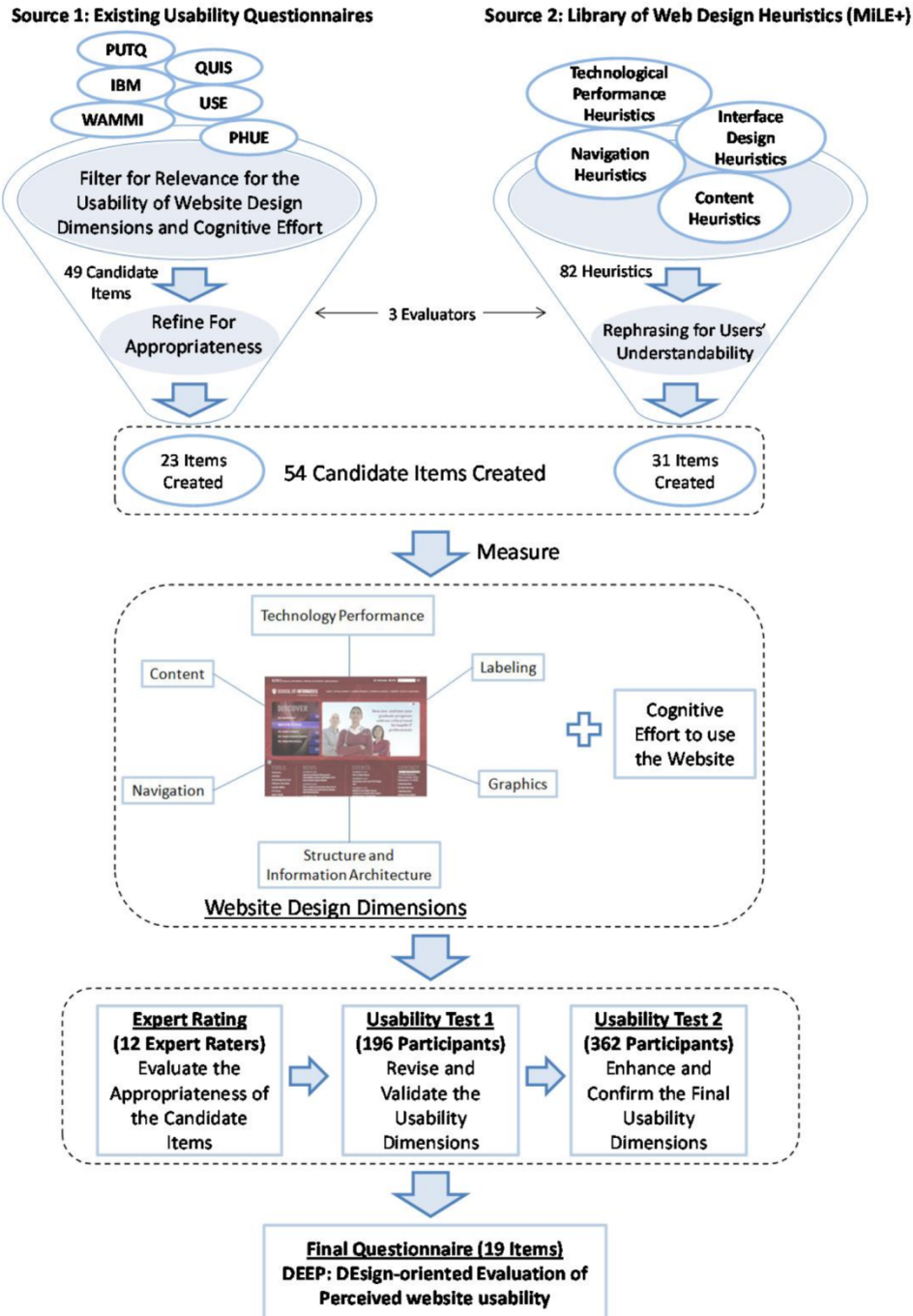


Figure 22. Procedure used to develop and validate the DEEP questionnaire: Identify candidate items → internal rating ($N = 12$) → online field test 1 ($N = 196$) → online field test 2 ($N = 362$) → final questionnaire.

4.1.5. Graphics

Rather than aesthetic attractiveness, this dimension focuses on the graphic designs that are critical for website usability, including the readability of the texts, organization of the

elements on a web page, and consistency of the web page layout (Baierova, Tate, & Hope, 2003; Bolchini & Garzotto, 2008; Bolchini & Paolini, 2004; Choi, Lee, & Kim, 2006; Park, 2007; Sharp, Rogers, & Preece, 2006; Tan & Wei, 2006). Good graphic designs could help users quickly become familiar with the entire website and easily locate the information they need, while poor graphics may lead users to irrelevant information (e.g., commercials) or confuse them with abrupt layout changes.

4.1.6. Technological Performance

The technological performance dimension concerns the possible inconvenience and errors that might occur while using a website. Examples of technological performance problems include long download time, unnecessary plug-ins, incompatibility of HTML code across different browsers, or 404 errors (i.e., page not found) (Bolchini & Garzotto, 2008; Bolchini & Paolini, 2004; Sharp, Rogers, & Preece, 2006). The scope of technological performance is not as clear as the other dimensions. It may stand for any error that occurs while browsing a website, including those caused by the browser and operating system, or only the errors caused by website itself (e.g., 404 errors). As shown by later parts of this chapter, potentially because of the complexity of pinpointing the causes of technological performance issues, I failed to capture the technological performance dimension using the selected measurement items.

Overall, this initial set of design dimensions represents the conceptual pillars that I used to extend the existing body of knowledge of perceived usability evaluation instruments. Based on this framework, the next section illustrates the methods used to create a new instrument that aims at measuring these dimensions from users' perspectives.

4.2. Research Methods and Techniques

The methodological procedure for creating DEEP is shown in Figure 22; a pool of candidate measurement items were created based on the proposed design dimensions. Then, the face validity of the candidate items was evaluated by student raters who are familiar with website usability. After a preliminary revision, the items were applied to the evaluation of actual information-intensive websites, which consisted of two rounds of online field tests. The aim of Online Field Test 1 was to explore and improve the initial design dimensions and measurement items under each dimension with the help of

exploratory and confirmatory factor analyses, whereas Online Field Test 2 aimed at further enhancing the questionnaire by only using a confirmatory factor analysis.

4.2.1. Creating Candidate Measurement Items Based on Design Dimensions

In addition to the six design dimensions proposed in 4.1, I added another dimension called cognitive effort as a gauge of the users' overall experiences with a website's usability. This dimension asked users whether the website as a whole was easy to learn and use (Bolchini & Garzotto, 2008; Han, 2000; Sharp, Rogers, & Preece, 2006). In this way, I avoided excluding from the questionnaire important usability concepts, such as learnability and memorability.

Two main sources were used to collect and create the candidate measurement items. The first source was the items of the existing usability questionnaires, in particular, PUTQ (Lin, Choong, & Salvendy, 1997), IBM Computer Usability Satisfaction Questionnaires (Lewis, 1995), WAMMI (Kirakowski, Claridge, & Whitehand, 1998), USE Questionnaire (Lund, 2001), PHUE (Perlman, 1997), and QUIS (Chin, Diehl, & Norman, 1998).

A Microsoft Excel data sheet was created with seven columns, each labeled as one of the six website design dimensions or as the cognitive effort dimension. Then, candidate measurement items were iteratively selected from the aforementioned questionnaires. Three evaluators were involved in this process. The first evaluator read the measurement items one-by-one. If an item was considered to be relevant to a given dimension, it was placed under the corresponding column in the data sheet. On finishing the initial sampling, the results were passed to the second evaluator who revised the list of items by either adding in new items or removing inappropriate items. Then, the sheet was passed to the third evaluator who also made revisions. After this process, the three evaluators worked together to discuss the results and make the final decision. In the end, 49 items were selected. The number of items chosen from each questionnaire and the distribution of these items across the seven dimensions are shown in Appendix E.

About half of the 49 items could not be directly used in the DEEP questionnaire as they were not directly related to website usability. For example, the items from QUIS can be used to evaluate the interface satisfaction of *any product*, and the IBM questionnaire is

suitable only for assessing the usability of *computer systems*. The true utility of these selected items was to provide inspirations in regard to what types of items were actually needed for measuring the usability of the selected design dimensions. Therefore, these items were either rephrased and revised or removed based on two criteria: an item should be relevant to website usability and the meaning of the item should be easily understood by website users who are not technically savvy. In the end, 23 candidate items were created (see Appendix F).

Another source for creating candidate measurement items was the MiLE+ heuristics library (Bolchini & Garzotto, 2008). The library provides up to 82 heuristics for usability experts performing analytical usability inspections of websites. However, the wording of these heuristics was too technical to be understood by ordinary users. Therefore, I transformed the relevant heuristics into perceived usability measurement items so that users who knew nothing about website usability could easily understand them. After several rounds of analysis, an additional set of 31 items was created (see Appendix F).

With the aforementioned sets of measurement items created, an initial questionnaire with 54 items was formed. Some of the items were *reversed*, meaning that I applied their negative meaning. For example, “Using this website made me feel tired.” These reversed items serve two purposes: prevent users from losing attention by solely working on positive items, and testing the negative side of a concept.

4.2.2. Performing Internal Ratings on the Candidate Items

Although the initial questionnaire went through several rounds of refinement, it was developed solely based on the three evaluators’ subjective judgments, which may lead to unexpected bias and reduce the effectiveness of the empirical validation process. Therefore, in order to enhance the reliability of the results, I pilot tested the appropriateness of the measurement items to the design dimensions on a larger sample of student raters. Twelve second-year graduate students were recruited from the Human–Computer Interaction program in the Indiana University School of Informatics. Each of these students had already studied and applied heuristic-based usability inspections in a graduate course called Usability and Evaluative Methods. Most of these students had full-time jobs in the informatics- or interface-design-related industry and had approximately three to five years of working experience in the field (in addition to the

educational experience in the HCI program). Although they may not be considered usability experts yet, they had a good sense of the importance of usability in everyday design practices.

The task for the raters was to rank the questionnaire items under each dimension based on the importance of the issue that an item touched on. For example, there are nine items under the content dimension. The raters were asked to rank them from one to nine, with one being the most important and nine being the least important. If the raters found that two items were equally important, they needed to force themselves to give different rankings to the items. In this way, a clearer pattern was obtained on the importance of the measurement items in order to evaluate each design dimension and the cognitive effort dimension. The results of the internal rating and how the questionnaire was further revised are summarized in the Results section.

4.2.3. Online Field Test 1: Revise and Validate Perceived Usability Dimensions

The next step was to apply the questionnaire to the evaluation of real-world websites (i.e., field tests) with a large sample of users. The field tests allowed me to collect enough instances of evidence in regard to whether the questionnaire could be used to accurately capture users' perceived usability and to run standard validation tests (e.g., exploratory factor analysis) with relatively high statistical power. The field tests were conducted online using the OFTEN environment (Chapter 3, section 3.1.4).

Website Selection. In order to thoroughly validate the questionnaire, it was important that the websites used in the field tests covered all of the design dimensions that the questionnaire was intended to assess. Therefore, instead of using simple and small-scale websites (i.e., websites with only a few web pages and a simple navigation paradigm), the selected websites had rich informational content, complex navigation systems and structures, and professional graphic designs. Moreover, two confounding variables needed to be controlled. First, the characteristics of websites from different domains are quite different. For example, with an entertainment website, users may care more about whether the audio and video on the site can be played correctly and fluently, whereas for a news website, users may focus more on the accuracy and timeliness of the information provided. Based on this concern, I found that university websites were a proper domain for the field tests. Compared to other types of websites, university

websites represented adequate design complexity and more balance across various design dimensions. Second, the reputation of a university was a possible confounding variable. People might be overly strict in judging the usability of an unknown university's website, but be tolerant of the usability problems found on a top-notch university's website. Therefore, in order to avoid this problem, the selected sites needed to be similar in terms of reputation level.

Another question that I had to address was how many websites were needed. According to the existing literature on psychometric instrument validation (Kirakowski, Claridge, & Whitehand, 1998; Lavie & Tractinsky, 2004; Sharp, Rogers, & Preece, 2006), an important indicator of the quality of a questionnaire is called predictive validity, which refers to whether a questionnaire can accurately measure the concept that it intends to capture. Therefore, in order to gauge my questionnaire using this indicator, at least two websites were needed: one with good usability and the other with poor usability. Using these two sites, I could see whether the evaluation results could accurately reflect the actual usability levels.

According to the requirements above, usability evaluations based on expert reviews (i.e., heuristic inspections) were performed on five candidate university websites, all of which were selected from the U.S. News & World Report (2010) national university rankings Tier 4. From these evaluations, two websites with significantly different usability ratings were identified. In order to protect the brand image of these institutions, I refrained from providing the actual names of the two university websites selected. In the rest of the chapter, I refer to them as "website with good usability" and "website with poor usability."

User Tasks. The field tests started by asking the participants to perform a number of tasks ($N = 8$) on a selected website. This activity allowed them to gain enough experience with every aspect of the website before they filled in the usability questionnaire. In order to make the experiences more natural, each of the user tasks was designed based on a scenario of an undergraduate student looking for opportunities in regard to applying for graduate school. As shown in Table 2, these tasks focused on various aspects of the university, including contact information, history, size, and tuition fees. The same tasks were used on both websites.

Table 2. Tasks used in online field test 1.

No.	Task
1	What is the mailing address of this university?
2	How many graduate students does this university have?
3	What year was this university established?
4	How many graduate programs does this university offer?
5	What is the minimum GPA you need to get into the graduate school at this university?
6	What is the price of one graduate credit hour for in-state residents?
7	What is the name for the school of business in this university?
8	What is the name of the basketball team of this university?

Participants. In order to be consistent with the user tasks, the target participants were *current undergraduate students*. Recruitment was conducted by sending 5,000 invitation letters (2,500 for each website) through e-mail to undergraduate students who took classes on one of the eight Indiana University– administered campuses during the spring semester of 2010. In the invitation letter, the purpose of the field test was briefly introduced and the URL to the customized OFTEN environment was provided. As an incentive, a report containing the results of the study was sent to all of the participants, and one participant was randomly selected to receive a \$100 prize.

Procedure. By clicking on the URL in the email, the participants were led to the OFTEN environment with features customized for this specific study. Four personal traits were collected at the beginning of the study: participant gender, age, education, and his/her intention to pursue a higher level education. Then, before performing the actual tasks, a warning message was show to tell the participants not to use the search box available on the site. The aim was to encourage the participants to navigate through the architecture and content of the website rather than simply *search* for information. The task performing interface was the same as the one shown in Figure 8 with the websites selected for this specific study embedded.

Upon finishing the tasks, the participants were asked to fill out the perceived usability questionnaire. Each item was measured on a 5-point Likert scale that ranged from strongly disagree to strongly agree included a not applicable option. This process was designed to take approximately 20 to 25 minutes.

Data Analysis. Both reliability and exploratory factor analyses were performed on the measurement items under each of the design dimensions in order to exclude irrelevant

or redundant items. Then, based on these analyses, I ran a confirmatory factor analysis using the LISREL 8.70 software in order to assess whether the revised dimensions were appropriate for measuring the concept of perceived usability as a whole. In the end, I created a revised questionnaire and came up with the strategies for the next round's online field test. The results of Online Field Test 1 can be found in section 4.3.2.

4.2.4. Online Field Test 2: Strengthening the Revised Perceived Usability Dimensions

Online Field Test 2 was basically a repeat of Online Field Test 1 (i.e., it contained similar procedures and the same user tasks) and aimed at enhancing and confirming the perceived usability dimensions established in the previous test. Only three changes were made.

First, the results from Online Field Test 1 indicated that the two websites used in the test failed to assess the questionnaire's predictive validity. The reason was that the usability of these two sites was not significantly different per se (the reason is elaborated upon in section 4.3.2). Therefore, from the list provided by Web (2009), I selected two other university websites, one with good usability and the other with poor usability according to the author. These two universities had similar rankings in the Forbes 2009 rankings of America's top 600 colleges (i.e., reflecting similar reputations). In addition, more in-depth heuristic inspections were performed in order to confirm that the usability levels of the two sites were indeed significantly different.

Second, in order to increase the robustness of the statistical model, I tried to recruit a larger sample of users. Therefore, more potential participants were contacted through e-mail (8,000 invitations were sent, 4,000 for each website), again using the Indiana University student directory.

Third, another important indicator for the quality of a psychometric measurement instrument is discriminant validity, which shows that the construct measured by the instrument is different from other similar constructs. Based on the existing literature, I found that visual aesthetics have a close relation to website usability (De Angeli, Sutcliffe, & Hartmann, 2006; Hartmann, Sutcliffe, & De Angeli, 2007; Hartmann, Sutcliffe, & De Angeli, 2008; Tractinsky, Katz, & Ikar, 2000). Therefore, in order to justify that my

questionnaire was trying to capture a construct different from visual aesthetics, I asked users to fill out a standardized website aesthetics questionnaire (Lavie & Tractinsky, 2004) at the end of Online Field Test 2. Although it was foreseeable that the results from these two questionnaires might be significantly correlated (De Angeli, Sutcliffe, & Hartmann, 2006; Hartmann, Sutcliffe, & De Angeli, 2007; Hartmann, Sutcliffe, & De Angeli, 2007; 2008; Tractinsky, Katz, & Ikar, 2000), according to Kline (1998), if the correlation was not higher than .85, I could still be confident that the questionnaire was assessing a different construct.

In next section, the results from the internal rating and the two rounds of online field tests are analyzed and the main findings from these results are presented.

4.3. Results

4.3.1. Results from the Internal Rating

The results for the internal rating are summarized in Table 3. Given that I asked the raters to give a ranking to the items under each dimension according to their importance, the smaller the number, the greater its importance. In addition, since the sample size was small (only 12 raters), only calculating the mean ranking scores might suggest bias due to the possible existence of extreme values. Therefore, both the mean and median scores were used to assess an item because the median was much more tolerant to outliers. If a great discrepancy existed between a mean and median score, it meant that the item was not necessarily reliable because there was too much variation among the raters' decisions.

Under the Content dimension, items 9 ("There were many irrelevant pictures on the website"; Mdn = 7.50, M = 7.25) and 10 ("There were many irrelevant videos on the website"; Mdn = 9.00, M = 8.08) ranked the lowest, which meant that displaying irrelevant pictures or videos was considered by the raters as the least important aspect in regard to the content usability of a website. Given that my aim was to measure the most relevant aspects for each design dimension, these two items were removed from the questionnaire. Under the Graphics dimension, Item 8 ("The fonts used on this website were strange") ranked the lowest (Mdn = 9.00, M = 8.58), which implied that the raters thought that the font style used on a website did not heavily influence the graphics usability. Therefore, I replaced it with a more general item: "The text on the website was

readable.” There were other items with low rankings, such as Item 8 (“Some web pages were too long”) in the Structure and Information Architecture dimension (Mdn = 8.00, M = 7.25). However, compared to the three items modified above, the rest were less obvious in terms of which actions (modify, replace, or remove) could be taken. So they were left unchanged.

Table 3. Internal rating of the perceived usability measurement items.

Dimension		Item No.										
		1	2	3	4	5	6	7	8	9	10	11
Content	Mdn	5.50	3.00	1.50	6.00	7.50	8.50	4.50	4.50	7.50	9.00	N/A
	M	4.92	3.33	2.58	5.75	6.67	6.83	5.08	4.50	7.25	8.08	
Structure and Information Architecture	Mdn	6.00	7.00	3.50	1.00	2.00	4.00	6.00	8.00	N/A		
	M	5.25	5.67	3.75	2.25	2.50	3.83	5.50	7.25			
Navigation	Mdn	4.00	4.50	7.00	2.00	6.50	6.00	8.00	4.00	6.50	5.50	N/A
	M	4.58	4.83	6.33	2.92	6.50	5.58	7.33	4.92	6.00	6.00	
Labeling	Mdn	2.00	3.00	4.50	4.00	6.00	7.00	2.50	N/A			
	M	2.17	3.17	4.08	4.08	5.58	6.42	2.50				
Cognitive Effort	Mdn	1.50	3.00	2.00	4.00					N/A		
	M	1.58	2.75	2.00	3.67							
Graphics	Mdn	8.50	4.50	4.00	6.00	8.00	6.00	5.00	9.00	4.50	3.00	4.00
	M	8.17	5.08	3.92	6.33	8.83	5.42	5.17	8.58	5.67	3.00	5.33
Technological Performance	Mdn	1.50	2.00	3.00	3.00					N/A		
	M	2.08	2.25	3.00	2.67							

Note. See Appendix F for the actual items.

Moreover, I noticed that Item 8 (“The text was concise and easy to read”) in the Content dimension was double-barreled, so it was split into two items: “The text was concise” and “The text was easy to read.” After these modifications, the revised questionnaire (see Appendix G) was used in Online Field Test 1.

4.3.2. Results from Online Field Test 1

Three hundred and seven responses were received (6.1% response rate) from Online Field Test 1. One hundred and ninety of the responses were valid (i.e., participants completed the entire test without skipping too many items or filling in fake answers). Among these valid responses, 94 were received from the test of the website with good usability (34 male, 60 female), while 102 were received from the website with poor usability (25 male, 77 female).

*The Perceived Content Dimension*². The perceived content dimension consisted of nine items (Table 4). The reliability analysis suggested that all of these items were reliably measuring the same concept (Cronbach's alpha = .820). Two factors were extracted from this dimension through exploratory factor analysis. The first factor was mainly about the understandability of the website content, which explained 36.94% of the total variation of the dimension. The second factor had to do with content errors, but only accounted for 20.48% of the total variation.

Table 4. Factor and reliability analyses of the perceived content dimension.

Item	Before Revision		After Revision
	Factor 1	Factor 2	Factor
2. The content (including text, pictures, audios, and videos etc.) was easy to understand.	.847		.838
1. The wording of the text was clear.	.824		.852
3. The content (including text, pictures, audios, and videos etc.) was what I expected.	.702		.660
9. The text was easy to read.	.688		.776
8. The text was concise.	.600		.699
7. The content looked trustworthy.	.551		
6. The content was up to date.	.467		
5. There were many grammatical errors and typos on the website.		.871	
4. The content (including text, pictures, audios, and videos etc.) did not contain errors or misspellings.		.771	
Total Variance Explained	36.94%	20.48%	59.08%
Cronbach's Alpha		.820	.819

In particular, I found that Item 5 ("There were many grammatical errors and typos on the website"), which had the highest loading in Factor 2, had a low correlation with the other items ($r = .299$). The reason might be that although grammatical errors and typos could give users a negative impression of the website, they may have little influence on the ease of use of the site. Therefore, the second factor (Items 4 and 5) was removed from the questionnaire. In addition, the questionnaire had 53 questions, which was too many compared to other existing usability questionnaires. In order to reduce its size, I decided to remove the items that had low factor loadings, such as, in this dimension, Items 6 ("The content was up to date") and 7 ("The content looked trustworthy"), which were removed due to factor loadings lower than .60. After these revisions, the following items were retained:

² In order to follow the reasoning of *perceived* usability, the term 'perceived' was added in front of each design dimension within the questionnaire. For example, perceived content, perceived navigation, and so forth.

- The content (including text, pictures, audios, and videos etc.) was easy to understand.
- The wording of the text was clear.
- The content (including text, pictures, audios, and videos etc.) was what I expected.
- The text was easy to read.
- The text was concise.

These items loaded to only one factor, which could explain 59.08% of the total variation of the dimension, while the Cronbach's Alpha stayed the same.

Table 5. Factor and reliability analyses of the perceived structure and information architecture dimension.

Item	Before Revision		After Revision
	Factor 1	Factor 2	Factor
15. Under each section of the website, the web pages were well organized.	.874		.858
14. The organization of the website was clear.	.846		.875
13. I could quickly get to know the structure of the website by skimming its homepage.	.784		.816
16. The links on each web page were well organized.	.780		.794
12. I could easily locate all related information in one section.	.749		.790
10. The long text on the website was well partitioned into sections.	.596		
11. Related web pages were scattered throughout the website.		.867	
17. Some web pages were too long.		.679	
Total Variance Explained	46.82%	17.50%	68.45%
Cronbach's Alpha	.848		.884

The Perceived Structure and Information Architecture Dimension. This dimension consisted of eight items (see Table 5; Cronbach's alpha = .848), which clustered into two factors. The first factor mainly concerned how well the content was organized in a website and whether users could quickly master the structure of the site, which explained 46.82% of the total variation of the dimension. The meaning of the second factor, however, was unclear.

In particular, both of the items in Factor 2, Items 11 ("Related web pages were scattered throughout the website") and 17 ("Some web pages were too long") were not highly correlated with the others ($r = .214$ and $.475$, respectively). For Item 17, contradictory evidence was found: Stone et al. (2005) suggested that web pages should not be too long, whereas Leech (2009) indicated that users do not mind scrolling down lengthy pages. Due to this debate, I decided to remove the item. Item 11 was also removed

because it might be difficult for users to define “related web pages.” In the end, the following items were retained:

- Under each section of the website, the web pages were well organized.
- The organization of the website was clear.
- I could quickly get to know the structure of the website by skimming its home page.
- The links on each web page were well organized.
- I could easily locate all related information in one section.

These items loaded to only one factor, which could explain 68.45% of the total variation of the dimension. The Cronbach’s alpha was improved to .884.

The Perceived Navigation Dimension. This dimension consisted of 10 items (see Table 6; Cronbach’s alpha = .882), which clustered into two factors. The first factor focused on navigating a website to find the desired information, which explained 47.73% of the total variation of the dimension. The second factor was related to the efficiency of navigating back to previous pages or sections, which explained only 16.25% of the total variation.

Table 6. Factor and reliability analyses of the perceived navigation dimension.

Item	Before Revision		After Revision
	Factor 1	Factor 2	Factor
21. It was easy to find the information I needed on the website.	.878		.892
26. This website helped me find what I was looking for.	.849		.864
19. The website provided enough guidance for me to navigate through the content.	.846		.866
27. Learning to find my way around this website was a problem.	.841		.862
25. It was difficult for me to move around this website.	.775		.784
18. It required only a few steps to accomplish the tasks.	.772		.805
22. Going from one section to another was easy on this website.	.661		
20. It was easy for me to return to previous pages.		.787	
24. Returning to higher-level pages was immediate.		.771	
23. Remembering where I was on this website was difficult.	.448	.497	
Total Variance Explained	47.73%	16.25%	71.61%
Cronbach’s Alpha	.882		.920

In the second factor, neither Item 20 (“It was easy for me to return to previous pages”) nor Item 24 (“Returning to higher level pages was immediate”) correlated highly with the other items ($r = .360$ and $.252$, respectively). The reason might be that even though a website itself does not provide links to go back to the previous pages, the users could still use the back button on the browser to go back. Therefore, this did not seem to be a notable usability problem for the users. Moreover, some of the items in Factor 1 already

cover the issue of *back navigation* in certain ways. For example, in item 21 (“It was easy to find the information I needed on the website”), during the process of “finding the needed information,” back navigation would be conducted for certain. Thus, I decided to remove these two items. With some additional modifications, the following items were retained:

- It was easy to find the information I needed on the website.
- This website helped me find what I was looking for.
- The website provided enough guidance for me to navigate through the content.
- Learning to find my way around this website was a problem.
- It was difficult for me to move around this website.
- It required only a few steps to accomplish the tasks.

These items loaded to only one factor, which could explain 71.61% of the total variation of the dimension. The Cronbach’s alpha was improved to .920.

The Perceived Labeling Dimension. This dimension consisted of 10 items (see Table 7; Cronbach’s alpha = .705), which clustered into two factors. The first factor concerned whether users could learn the meaning of the labels quickly, which explained 43.87% of the total variation of the dimension. The second factor mainly assessed the understandability of the captions for pictures and videos, which explained 23.86% of the total variation.

Table 7. Factor and reliability analyses of the perceived labeling dimension.

Item	Before Revision		After Revision
	Factor 1	Factor 2	Factor
29. When I skimmed through the labels I confidently knew where to click.	.834		.838
34. I got what I expected when I clicked on things on this website.	.833		.842
28. The meanings of the main navigation labels were straightforward.	.819		.818
30. The links on the website were clear.	.769		.800
31. The headings of the web pages were easy to understand.	.634	.417	
32. The captions of the pictures were clear.		.873	
33. The captions of the videos were clear.		.850	
Total Variance Explained	43.87%	23.86%	68.05%
Cronbach’s Alpha	.705		.843

In the second factor, Items 32 (“The captions of the pictures were clear”) and 33 (“The captions of the videos were clear”) had the highest factor loadings, but did not correlate well with the other items ($r = .397$ and $.205$, respectively). This indicated that although

the captions for multimedia objects were critical for website accessibility of visually impaired users (W3C, 2008), they did not greatly affect sighted users' usability experiences. This was because the sighted users could simply view the pictures or videos to grasp their meanings. Therefore, these two items were removed from the questionnaire. With some additional modifications, the following questions were retained:

- When I skimmed through the labels I confidently knew where to click.
- I got what I expected when I clicked on things on this website.
- The meanings of the main navigation labels were straightforward.
- The links on the website were clear.

These items loaded to one factor, which could explain 68.05% of the total variance of the dimension. The Cronbach's alpha was improved to .843.

Table 8. Factor and reliability analyses of the perceived cognitive effort.

Item	Before Revision	After Revision
	Factor	Factor
37. I learned to use this website quickly.	.882	.905
35. Using this website was effortless.	.843	.874
36. Using this website made me feel tired.	.837	.838
38. The web pages were dense.	.545	
Total Variance Explained	62.17%	76.19%
Cronbach's Alpha	.789	.840

The Perceived Cognitive Effort Dimension. This dimension consisted of four items (see Table 8; Cronbach's alpha = .789), which all loaded to one factor. However, the factor loading for item 38 ("The web pages were dense") was relatively low (.545), and its correlation with the other items was only .365. A possible reason could be that the meaning of the word "dense" was vague. Therefore, I decided to remove the item. After this slight modification, following items were retained:

- I learned to use this website quickly.
- Using this website was effortless.
- Using this website made me feel tired.

The total variance explained increased from 62.17% to 76.19% and the Cronbach's alpha increased to .840.

The Perceived Graphics Dimension. This dimension consisted of 11 items (see Table 9; Cronbach's $\alpha = .826$). An exploratory factory analysis clustered them into four factors: the readability of texts, visual guidance through highlighting and colors, consistency of page layout, and misuse of colors, which respectively accounted for 22.25%, 20.26%, 15.44%, and 13.90% of the total variance of the dimension. In order to ensure that Factors 1 and 2 were mutually exclusive constructs, I decided to remove items 47 ("The home page of this website was pleasant"), 42 ("The pages on this website were pleasant to look at"), and 46 ("The colors of the text made it easy to read") because they loaded on both Factors 1 and 2. Factors 3 and 4 only consisted of two items, which was considered be not reliable and to need further improvements (in general, at least three items are needed for each factor). Different from the other dimensions, all four factors seemed to be important to website usability. Therefore, they were all kept for the confirmatory factor analysis.

Table 9. Factor and reliability analyses of the perceived graphics dimension (original).

Item	Before Revision			
	Factor 1	Factor 2	Factor 3	Factor 4
45. The text on the website was readable.	.893			
44. The text on the website was big enough to read.	.885			
47. The home page of this website was pleasant.	.547	.497		
40. The colors helped me to distinguish different sections of the website.		.826		
41. The highlighted areas of a page helped me locate the information I needed.		.778		
42. The pages on this website were pleasant to look at.	.477	.536		
46. The colors of the text made it easy to read.	.450	.454		
49. I noticed abrupt changes in the layout of the pages.			.880	
48. The layout of the pages throughout the website was consistent.			.827	
39. The colors used on the website made me feel uncomfortable.				.794
43. This website had some annoying colors.				.747
Total Variance Explained	22.25%	20.26	15.44	13.90
Cronbach's Alpha		.826		

Table 10. Factor and reliability analyses of the perceived graphics dimension (revised).

Item	After Revision			
	Factor 1	Factor 2	Factor 3	Factor 4
44. The text on the website was big enough to read.	.917			
45. The text on the website was readable.	.906			
49. I noticed abrupt changes in the layout of the pages.		.894		
48. The layout of the pages throughout the website was consistent.		.828		
40. The colors helped me to distinguish different sections of the website.			.871	
41. The highlighted areas of a page helped me locate the information I needed.			.814	
39. The colors used on the website made me feel uncomfortable.				.880
43. This website had some annoying colors.				.700
Total Variance Explained	22.37%	20.53%	19.75%	16.68%
Cronbach's Alpha		.720		

Table 10 shows the revised perceived graphics dimension. The four factors respectively explained 22.37%, 20.53%, 19.75%, and 16.68% of the total variance of the dimension. Cronbach's alpha dropped to .720 but was still higher than the .7 threshold.

The Perceived Technological Performance Dimension. The reliability analysis for this dimension (Table 11) was not satisfactory: Cronbach's alpha = .322 (far below the .7 threshold). The correlations among the measurement items were quite low and it was impossible to substantially improve the Cronbach's alpha by removing any of the items. Moreover, as discussed in section 4.1.6, the notion of technological performance covers too many aspects that do not necessarily have to do with website design. Therefore, I decided to eliminate this entire dimension from the questionnaire.

Table 11. Reliability analysis of the perceived technological performance dimension.

Item	Corrected Item-Total Correlation	Alpha if Item Deleted
50. The pages took too long to load.	.0135	.3871
51. I needed to install plug-ins to view some of the content.	.1275	.3154
52. When I made a mistake, the website promptly helped me to recover.	.2112	.2159
53. Multimedia (audio and video) functioned well on this website.	.3177	.0185
Cronbach's Alpha	.322	

Based on the improvements summarized above, nine perceived usability dimensions were extracted: perceived content, structure and information architecture, navigation, labeling, cognitive effort, and the four perceived graphics factors. Before conducting a confirmatory factor analysis, I first checked the correlations among these dimensions in order to understand whether they were properly correlated with each other and whether the correlations were too high. This could offer me some preliminary insights regarding which dimension(s) should be removed or merged.

Correlations Among the Revised Dimensions. The full correlation matrix of the improved usability dimensions can be found in Appendix H. All of the correlations were statistically significant ($p < .01$), meaning that they were measuring the same concept (i.e., perceived usability). The correlations among the perceived content and the four perceived graphics dimensions were within a proper range ($r < .536$) and their correlations with the other four dimensions were also not overly high ($r < .67$), indicating that they were measuring different sub-constructs. In contrast, the correlations among

the perceived structure and information architecture, navigation, labeling, and cognitive effort dimensions were relatively high ($.75 < r < .90$), which implied that some of these dimensions could be merged. However, instead of making hasty changes based on the correlations, the decisions were made using a confirmatory factor analysis.

Confirmatory Factor Analysis. A confirmatory factor analysis was performed (see Table 12) using the LISREL 8.70 software on the nine perceived usability dimensions extracted from the exploratory factor analysis. The initial model did not show enough goodness of fit: $\chi^2 = 844.88$, goodness-of-fit index (GFI) = .78, adjusted goodness-of-fit index (AGFI) = .75, normal fit index (NFI) = .95, comparative fit index (CFI) = .98, root mean square residual (RMR) = .066, and root mean square error of approximation (RMSEA) = .000. A favorable GFI and AGFI should be greater than .90, and the RMR should be lower than .05. In addition, the factor loadings for each dimension or item should not be too low (at least higher than .6).

Table 12. Structural coefficients for the perceived usability dimensions.

Item	Perceived Usability Dimension								
	Cont	Strut	Nav	Lab	Cog	G1	G2	G3	G4
1. The wording of the text was clear.	.83								
2. The content (including text, pictures, audios, and videos etc.) was easy to understand.	.81								
9. The text was easy to read.	.69								
8. The text was concise.	.59								
3. The content (including text, pictures, audios, and videos etc.) was what I expected.	.58								
14. The organization of the website was clear.		.87							
15. Under each section of the website, the web pages were well organized.		.78							
13. I could quickly get to know the structure of the website by skimming its homepage.		.78							
12. I could easily locate all related information in one section.		.74							
16. The links on each web page were well organized.		.71							
21. It was easy to find the information I needed on the website.			.88						
19. The website provided enough guidance for me to navigate through the content.			.84						
26. This website helped me find what I was looking for.			.83						
27. Learning to find my way around this website was a problem.			.82						
18. It required only a few steps to accomplish the tasks.			.76						
25. It was difficult for me to move around this website.			.75						
34. I got what I expected when I clicked on things on this website.				.84					

29. When I skimmed through the labels I confidently knew where to click.										.75
28. The meanings of the main navigation labels were straightforward.										.74
30. The links on the website were clear.										.69
37. I learned to use this website quickly.										.88
35. Using this website was effortless.										.80
36. Using this website made me feel tired.										.73
45. The text on the website was readable.										.90
44. The text on the website was big enough to read.										.83
48. The layout of the pages throughout the website was consistent.										.96
49. I noticed abrupt changes in the layout of the pages.										.61
41. The highlighted areas of a page helped me locate the information I needed.										.85
40. The colors helped me to distinguish different sections of the website.										.61
43. This website had some annoying colors.										.76
39. The colors used on the website made me feel uncomfortable.										.44
Overall Factor Loading to Perceived Usability	.71	.96	.95	.99	.98	.46	.58	.63	.39	
Cronbach's Alpha	.819	.884	.920	.843	.840	.839	.736	.679	.501	
Model	χ^2	df	GFI	AGFI	NFI	CFI	RMR	RMSEA		
	844.88	425	.78	.75	.95	.98	.066	.000		

Note: Cont = Content; Strut = Structure and Information Architecture; Nav = Navigation; Cog = Cognitive Effort; Lay = Layout Consistency = Graphics_2; Vis = Visual Guidance = Graphics_3; GFI = goodness-of-fit index; AGFI = adjusted goodness-of-fit index; NFI = normal fit index; CFI = comparative fit index; RMR = root mean square residual; RMSEA = root mean square error of approximation.

In order to improve this initial model, I first noticed that the overall factor loadings for Graphics_1 and Graphics_4 were quite low (.46 and .39, respectively), which meant that they were not highly qualified dimensions for measuring the concept of perceived usability. Hence, I eliminated them from the model.

Second, in each dimension, some of the items had a factor loading lower than .70, which was considered not satisfactory; however, instead of just removing all of them right away, I tried to not influence too much of the reliability of a dimension (Cronbach's alpha should be greater than .70). In case the overall Cronbach's alpha would be decreased too much by deleting an item, I would rather keep it in the model. With these considerations in mind, only Items 3 ("The content (including text, pictures, audios, and videos etc.) was what I expected") and 30 ("The links on the website were clear") were eliminated.

Table 13. Structural coefficients for the perceived usability dimensions (after revision).

Item	*Perceived Usability Dimension					
	Cont	Strut	Nav	Cog	Lay	Vis
1. The wording of the text was clear.	.85					
2. The content (including text, pictures, audios, and videos etc.) was easy to understand.	.78					
9. The text was easy to read.	.70					
8. The text was concise.	.60					
14. The organization of the website was clear.		.91				
13. I could quickly get to know the structure of the website by skimming its homepage.		.80				
15. Under each section of the website, the web pages were well organized.		.75				
21. It was easy to find the information I needed on the website.			.88			
34. I got what I expected when I clicked on things on this website.			.85			
26. This website helped me find what I was looking for.			.82			
27. Learning to find my way around this website was a problem.			.81			
37. I learned to use this website quickly.				.89		
35. Using this website was effortless.				.80		
36. Using this website made me feel tired.				.73		
48. The layout of the pages throughout the website was consistent.					.95	
49. I noticed abrupt changes in the layout of the pages.					.62	
41. The highlighted areas of a page helped me locate the information I needed.						.87
40. The colors helped me to distinguish different sections of the website.						.60
Overall Factor Loading to Perceived Usability	.68	.93	.97	.99	.54	.60
Cronbach's Alpha	.819	.854	.840	.843	.736	.679
Model	χ^2	df	GFI	AGFI	NFI	CFI
	241.65	129	.88	.84	.96	.98
	RMR	RMSEA				
	.061	.019				

Note: Cont = Content; Strut = Structure and Information Architecture; Nav = Navigation; Cog = Cognitive Effort; Lay = Layout Consistency = Graphics_2; Vis = Visual Guidance = Graphics_3; GFI = goodness-of-fit index; AGFI = adjusted goodness-of-fit index; NFI = normal fit index; CFI = comparative fit index; RMR = root mean square residual; RMSEA = root mean square error of approximation.

Third, as stated in the previous section, the correlations among structure and information architecture, navigation, labeling, and cognitive effort were pretty high ($.75 < r < .90$), which suggested that they could possibly be merged into one dimension; however, when I tried to conduct the merge, I found that the goodness-of-fit of the model decreased in most cases, except when only the navigation and labeling dimension were combined. This implied that the structure and information architecture and the cognitive effort dimensions should be kept as separate dimensions. However, in the users' opinions, the navigation system and labels that a website uses offered them the same kind of experiences because users navigate a website with the guidance of the labels.

Therefore, they were merged into one dimension, while the name of the dimension remained unchanged as perceived navigation.

Finally, in order to reduce the size of the questionnaire, several items with low factor loadings or redundant meanings were removed (on the basis of not affecting reliability). The results of the improved model are summarized in Table 13.

After the revision, the model was greatly improved: $\chi^2 = 241.65$, GFI = .88, AGFI = .84, NFI = .96, CFI = .98, RMR = .061, and RMSEA = .019. In this model, the concept of perceived usability was operationalized into six dimensions: content, structure and information architecture, navigation, cognitive effort, Graphics_2, and Graphics_3. According to the meanings of the corresponding measurement items, Graphics_2 was renamed *Layout Consistency* and Graphics_3 was renamed *Visual Guidance*.

Nevertheless, the model still needed further improvement as there were still items in the perceived content dimension which had low factor loadings and the layout consistency and visual guidance dimensions consisted of only two items, which could weaken the stability of the model. In order to address these issues, new candidate measurement items were added to these dimensions for the next round's online field test (see Appendix J for the questionnaire used for Online Field Test 2). These new items were created either by improving the wording of the existing low-factor-loading items or by adding the missing aspects which were supposed to be covered by a dimension.

Validation and Implications for Online Field Test 2. Another key quality for a psychometric instrument is its *validity*, which concerns whether the evaluation instrument I developed (i.e., the perceived usability questionnaire) could precisely reflect the reality (i.e., website usability). Three types of validity were analyzed on the current questionnaire, which included external validity, convergent validity, and predictive validity. During the process of the validity analysis, the implications for Online Field Test 2 were proposed.

External validity refers to whether an assumed relation could be justified. In this study, it was expected that higher task success rates would yield better perceived usability. Therefore, the external validity of the questionnaire could be calibrated by checking

whether there was a significant correlation between the users' success rates on the eight tasks that they performed (i.e., performance-based usability) and the perceived website usability, which was measured by the questionnaire. As shown in Appendix I, a two-tailed correlation test indicated that there was a significant correlation between task success rate and the overall perceived usability (calculated by averaging all of the usability dimensions; $r = .368$, $p < .01$). Particularly, the task success rate was significantly correlated with five of the six perceived usability dimensions (except for the layout consistency dimension). Therefore, the external validity of the questionnaire was considered satisfactory.

Convergent validity has to do with whether the measurement items and sub-constructs in a questionnaire measure the same higher-level concept. Based on the foregoing reliability analyses (Cronbach's Alpha) within each perceived usability dimension and the significant correlations among these dimensions (Appendix I), I could conclude that the convergent validity of the current questionnaire was sufficient.

Predictive validity focuses on whether a questionnaire can accurately measure the concept that it intends to measure. In this study, it refers to whether the perceived usability questionnaire can accurately predict the usability of a website. Unfortunately, however, I found that the poor-usability website did not yield a significantly lower success rate. The average task success rate for the good-usability website was 76.2% (SD = .187), while it was 77.9% (SD = .181) for the poor-usability website ($t = -.663$, $p > .05$). Therefore, the websites used in Online Field Test 1 were unlikely to be good platforms by which to assess the predictive validity of the questionnaire (i.e., the questionnaire could not predict which website had better usability because the websites being tested were at the same usability level). In fact, this problem was reflected by the questionnaire: the average perceived usability score for the good-usability website was 3.25/5 (SD = .77) and 3.36/5 (SD = .73) for the poor-usability site, which did not show a significant difference ($t = -.956$, $p > .05$). This problem could be caused by insufficient heuristic inspections or the personal bias of the three evaluators. Therefore, as introduced in the Research Methods and Techniques section, in Online Field Test 2, I selected two other websites based on other researchers' recommendations (Fishman, 2009) and more in-depth heuristic inspections.

4.3.3. Results from Online Field Test 2

Four hundred and seventy-six responses were collected from Online Field Test 2 (5.95% response rate). Three hundred and sixty-two of these responses were valid. Among the valid responses, 187 were collected from the test of the website with good usability (59 male, 128 female), while 175 were from the website with poor usability (57 male, 118 female).

Table 14. Structural coefficients for the perceived usability dimensions (final).

Item	*Perceived usability Dimension					
	Cont	Strut	Nav	Cog	Lay	Vis
Perceived Content						
The text was useful.	.85					
The text was relevant.	.84					
The content (including text, pictures, audios, and videos etc.) was easy to understand.	.72					
The wording of the text was clear.	.71					
Perceived Structure and Information Architecture						
The organization of the website was clear.		.90				
I could quickly get to know the structure of the website by skimming its homepage.		.83				
Under each section of the website, the web pages were well organized.		.83				
Perceived Navigation						
It was easy to find the information I needed on the website.			.91			
This website helped me find what I was looking for.			.88			
I got what I expected when I clicked on things on this website.			.87			
Perceived Cognitive Effort						
I learned to use this website quickly.				.88		
Using this website was effortless.				.82		
Using this website made me feel tired.				.71		
Perceived Layout Consistency						
The layout of the pages throughout the website was consistent.					.90	
The layout under each section of the website was consistent.					.85	
I noticed abrupt changes in the layout of the pages.					.76	
Perceived Visual Guidance						
The highlighted areas of a page helped me locate the information I needed.						.89
I got to know the content of a page by skimming the highlighted areas.						.83
The colors helped me to distinguish different sections of the website.						.67
Overall Factor Loading to Perceived Usability	.79	.97	.97	.97	.67	.79
Cronbach's Alpha	.859	.887	.916	.842	.832	.872
Model	χ^2	df	GFI	AGFI	NFI	CFI
	337.78	146	.91	.88	.98	.99
	RMR	RMSEA				
	.044	.022				

Note: Cont = Content; Strut = Structure and Information Architecture; Nav = Navigation; Cog = Cognitive Effort; Lay = Layout Consistency = Graphics_2; Vis = Visual Guidance = Graphics_3; GFI = goodness-of-fit index; AGFI = adjusted goodness-of-fit index; NFI = normal fit index; CFI = comparative fit index; RMR = root mean square residual; RMSEA = root mean square error of approximation.

Since the aim of Online Field Test 2 was to enhance and confirm the usability dimensions established in Online Field Test 1, I did not conduct another round of exploratory factor analysis. Instead, a confirmatory factor analysis was directly performed on the new data collected. With several minor adjustments – removing items with low factor loading from the perceived content, layout consistency, and visual guidance dimensions – the final questionnaire was established (see Table 14 or Appendix K for the actual questionnaire).

Goodness-of-fit. The final model with six perceived usability dimensions (see Table 13) showed a good fit to the data: $\chi^2(146) = 337.78$, GFI = .91, NFI = .98, and CFI = .99. AGFI (= .88) was slightly below .90, which was still acceptable. RMR and RMSEA were below the .05 standard. Under each dimension, all items had factor loadings higher than .60 and most were higher than .70, which indicated that they were good instruments for gauging the usability of each of the website design dimensions and users' cognitive efforts in using a website.

Reliability. The Cronbach's alpha for the entire questionnaire was .954 (>.7), which suggested that all of the items in the questionnaire were measuring the same concept (i.e., perceived usability). For each of the design-based dimensions (i.e. perceived content, structure and information architecture, navigation, layout consistency, and visual guidance), the Cronbach's alpha was .859, .887, .916, .832, and .872, respectively. The Cronbach's alpha for the perceived cognitive effort dimension was .842. Therefore, each dimension could also be reliably captured by the corresponding measurement items.

Validity. Five types of validity were evaluated on the final questionnaire: predictive validity, external validity, convergent validity, discriminant validity, and face validity.

In this round, the two testing websites yielded significantly different task success rates, which allowed me to better analyze the questionnaire's predictive validity. The task success rate for the good-usability website was 84.49% (SD = .179), while it was 59.93% (SD = .208) for the poor-usability site ($t = 12.0$, $p < .01$). Based on this, I found that the questionnaire had accurate predictions. The average perceived usability score of the good-usability website was 3.98/5 (SD = .55), which was significantly higher than that of

the poor-usability website 2.85/5 (SD = .63). Therefore, the predictive validity of the questionnaire could be considered satisfactory.

Two-tailed correlation tests were performed between the perceived usability dimensions and the average task success rates (Appendix L). The results showed that each of the perceived usability dimensions and the overall average perceived usability scores had positive and significant correlations with task success rates at the .01 level, which suggested a good external validity of the questionnaire.

Good convergent validity could already be justified through the reliability analyses. The significant correlation among the usability dimensions was another sign of high convergent validity (Appendix L).

Discriminant validity was assessed by checking whether the perceived usability questionnaire was measuring a concept different from other closely-related concepts. As mentioned in section 4.2.4, the concept that I needed to pay special attention to was the aesthetics of a website. In particular, the standardized aesthetics questionnaire that I used operationalized aesthetics into two dimensions: classical and expressive aesthetics. As expected, the two-tailed correlation analysis (Appendix M) showed that each of the perceived usability dimensions and the overall perceived usability significantly correlated with the classical, expressive, and overall aesthetics. However, none of these correlations were higher than .85. As suggested by Kline (1998), this was a good indication that the questionnaire had satisfactory discriminant validity. As for the discriminant validity among the perceived usability dimensions, although they were significantly correlated (Appendix M), none of these correlations were higher than .85. Therefore, the discriminant validity within the questionnaire was also good.

Face validity concerned whether the perceived usability dimensions I elicited would make sense to website usability professionals in general. The face validity of the questionnaire was first ensured by carefully creating the candidate measurement items (section 4.2.1) and then by having HCI graduate students rate the items (section 4.2.2). Second, in the final questionnaire, as stated in the review of the background, theories already existed that support the perceived content, structure and information architecture, navigation, and cognitive effort dimensions. As for the two dimensions extracted from the

website's graphic design, perceived layout consistency was related to one of the 10 usability heuristics devised by Nielsen (1993), while visual guidance was emphasized in the usability literature in terms of web page scannability (Stone et al., 2005).

Based on the above analysis, the final DEEP questionnaire could be considered a highly qualified instrument for measuring the perceived usability of five website design dimensions as well as perceived cognitive effort, in terms of both good reliability and validity.

4.4. Discussion

4.4.1. Perception of Labeling Versus Navigation

Based on the factor analysis results from Online Field Test 1, I made changes to the initial usability dimensions. In particular, I found that the measurement items under the perceived navigation and labeling dimensions clustered into one factor. This revealed that the users had different mindsets from those of the designers in regard to conceptualizing the two dimensions. Designers consider navigation to be the accessing mechanism of a website, whereas labeling focuses on naming the navigation links and page elements. However, users may never think of the two concepts separately. For them, they just read the label of a navigation link, click on it, and open the underlying page. Both the link and its label are accessing mechanisms, not separable features. For this reason, I combined the items relating to labeling and navigation into one dimension, but kept its name as Navigation. This strongly suggests that both navigation and labeling issues account for what users perceive as being able to find information based upon clues given to them within the site, regardless of analytical design concerns that those issues actually stem from. This finding was not extracted from the heuristic analysis that the evaluators conducted initially. It was also a good example of how users perceive usability differently and, one could argue, less technically (Ssemugabi & de Villiers, 2007) than usability experts.

4.4.2. Perceived Technical Performance

The technical performance dimension was eliminated entirely from the questionnaire due to the inability of the corresponding measurement items to reliably capture the intended concept (Cronbach's $\alpha < .7$). Moreover, the results did not offer any clues regarding what types of items would be suitable to measure this dimension. One potential reason

for this might be that the users did not think that the technical issues specified by the measurement items (Appendix G) could be put into a same usability category (i.e., technical performance). For example, a page downloading problem might have nothing in common with a plug-in error for general website users. Usability experts might consider both problems caused by technological obstacles or dysfunctions, but this categorization criterion was unlikely to be understood by non-technical-savvy users. Another reason could be that users separate some of the technical performance issues of a website from what they would consider its usability. For example, they might think of a dysfunctional plug-in as a third-party problem (e.g., Adobe flash player), rather than a mistake made by the site creators. This could be why the measurement items were so loosely correlated. No matter which of the above reasons was the actual reason, it is obvious that general website users cannot perceive all technical performance issues as a single conceptual dimension and, therefore, further research is needed to understand whether users would attribute a technical problem to a mistake made by the website designers or to a third-party service provider, such as an Internet service or plug-in provider.

4.4.3. Elements of Graphics Usability

Another major modification that I made to the design dimensions was to split the perceived graphics dimension into two dimensions: perceived layout consistency and perceived visual guidance. This indicates that rather than perceiving the abstract concept of 'graphics,' users assessed the usability of website graphics by something more concrete, such as whether there was an abrupt layout change between pages or whether the colors and highlights on the page could help them locate the information they needed. This finding is in line with other studies on the perception and requirements for usable graphics on the web. As for the perceived layout consistency dimension, users desired a consistent layout (Sutcliffe, 2002), which helped to keep them focused on the website content, while not allowing them to stray from the intended path. As for the perceived visual guidance dimension, graphics helped show users where they were within a site and helped them to navigate through the site's content (Fang & Holsapple, 2007).

4.4.4. Minimizing the Number of Items

I also reduced the number of measurement items in order to avoid user fatigue. Using the techniques described throughout this chapter, I removed items that had low factor loadings – which translated to their relatively low ability to explain the variances of a dimension – in order to limit the number of items, while not sacrificing the reliability of the instrument. Compared to the questionnaires I used to create the initial items, QUIS (27 items), PUTQ (100 items), CSUQ (19 items), USE (30 items), WAMMI (20 items), and PHUE (13 items), the DEEP questionnaire is of a reasonable length (19 items). With this size, I expect that users can reasonably complete all of the answers in approximately five to 10 minutes.

4.4.5. Guidance for Using and Interpreting DEEP: The Designer's Perspective

DEEP serves as a precise diagnostic tool to identify usability breakdowns in the design of web systems, as emerging in the user experience. In order to make the best use of DEEP, a simple numeric scale can be associated with the questionnaire items. For example, we recommend using one to five values with three for Neutral or N/A (or using a 7-point scale with four as the Neutral or N/A point) associated to the Likert scale of each DEEP item. In this way, it is easy to calculate the average perceived usability scores at three levels: an overall score of perceived usability, a score for each design dimension, and a score for each item.

How can designers interpret these resulting scores? I suggest focusing on the most analytical level of evaluation (item level) within each design dimension, as it can provide the most useful insight into the aspects of the design that causes usability issues in the user experience.

Table 15. Guiding the interpretation of DEEP results – examples.

Item	What to do in Case of a Score ≤ 3 <i>(Examples of Actions to be Taken and Resources to be Used)</i>
Perceived Content	
The wording of the text was clear	Seek ways to simplify the style and language of the text content. <i>See research-based guidelines for usable design:</i> http://usability.gov/guidelines/guidelines_book.pdf

<p>The content (including text, pictures, audios, and videos etc.) was easy to understand</p>	<p>Chapters 11 and 15.</p> <p>Seek ways to simplify the style and language of the content (including text, pictures, audio, and videos, etc.).</p> <p><i>See research-based guidelines for usable design:</i> http://usability.gov/guidelines/guidelines_book.pdf</p>
<p>The text was useful</p>	<p>Chapters 11 and 15.</p> <p>Seek ways to identify and cut redundancies in the text content. Refocus on the key content to support the tasks of the target users.</p> <p><i>See research-based guidelines for usable design:</i> http://usability.gov/guidelines/guidelines_book.pdf</p>
<p>The text was relevant</p>	<p>Chapters 11 and 15.</p> <p>Seek ways to identify and cut redundancies in the text content. Refocus on the key content to support the tasks of the target users.</p> <p><i>See research-based guidelines for usable design:</i> http://usability.gov/guidelines/guidelines_book.pdf</p>
<p>The text was relevant</p>	<p>Chapters 11 and 15.</p> <p>Seek ways to identify and cut redundancies in the text content. Refocus on the key content to support the tasks of the target users.</p> <p><i>See research-based guidelines for usable design:</i> http://usability.gov/guidelines/guidelines_book.pdf</p>
<p>The text was relevant</p>	<p>Chapters 11 and 15.</p> <p>Seek ways to identify and cut redundancies in the text content. Refocus on the key content to support the tasks of the target users.</p> <p><i>See research-based guidelines for usable design:</i> http://usability.gov/guidelines/guidelines_book.pdf</p>
<p>Perceived Structure and Information Architecture</p>	
<p>I could quickly get to know the structure of the website by skimming its homepage</p>	<p>Re-conceive and re-design the logical structure of the homepage by following the many heuristics and guidelines available on the topic.</p> <p><i>See research-based guidelines for usable design:</i> http://usability.gov/guidelines/guidelines_book.pdf</p>
<p>The organization of the website was clear</p>	<p>Chapter 5.</p> <p>Re-conceive and re-design the overall information architecture, in terms of basic access structures, navigation paths, and orientation mechanisms.</p> <p><i>See research-based guidelines for usable design:</i> http://usability.gov/guidelines/guidelines_book.pdf</p>
<p>Under each section of the website, the web pages were well organized</p>	<p>Chapters 12 and 16.</p> <p>Re-conceive and re-design the organization of each section of the site, providing consistent structures and orientation mechanisms across sections.</p> <p><i>See research-based guidelines for usable design:</i> http://usability.gov/guidelines/guidelines_book.pdf</p>
<p>Perceived Navigation</p>	
<p>It was easy to find the information I needed on the website</p>	<p>Chapters 12 and 16.</p> <p>Refocus overall navigation structure, starting from the homepage, on key organizational schemes: provide access to information by tasks, target audience, and topical categories. If feasible, provide an efficient and accurate search function within the website.</p> <p><i>See research-based guidelines for usable design:</i> http://usability.gov/guidelines/guidelines_book.pdf</p>

This website helped me find what I was looking for

Chapters 7 and 17.

Refocus overall navigation structure, starting from the homepage, on key organizational schemes: provide access to information by tasks, target audience, and topical categories. If feasible, provide an efficient and accurate search function within the website.

See research-based guidelines for usable design:
http://usability.gov/guidelines/guidelines_book.pdf

Chapters 7 and 17.

I got what I expected when I clicked on things on this website

Re-check the clarity and transparency of the link labels with respect to the background knowledge of your target audience.

See research-based guidelines for usable design:
http://usability.gov/guidelines/guidelines_book.pdf

Chapters 9 and 10.

Perceived Cognitive Effort

Using this website was effortless

See research-based guidelines for usable design:

Te'eni, D., Carey, J., & Zhang, P. (2006). HCI: Developing Effective Organizational Information. John Wiley & Sons, Inc.

Chapter 8 – 3.2 User Control and Feedback Guidelines

*Using this website made me feel tired
(Important: This is a reversed item.
Please give inverted ratings: 5 = strongly disagree, 4 = disagree, 3 = neutral, 2 = agree, 1 = strongly agree, 0 = Not Applicable)

See research-based guidelines for usable design:

Te'eni, D., Carey, J., & Zhang, P. (2006). HCI: Developing Effective Organizational Information. John Wiley & Sons, Inc.

Chapter 8 – 3.2 User Control and Feedback Guidelines

I learned to use this website quickly

See research-based guidelines for usable design:

Linja-aho, M. (2005). Evaluating and Improving the Learnability of a Building Modeling System. Helsinki University of Technology.

Perceived Layout Consistency

The layout of the pages throughout the website was consistent

See research-based guidelines for usable design:
http://usability.gov/guidelines/guidelines_book.pdf

Chapter 6

* I noticed abrupt changes in the layout of the pages
(Important: This is a reversed item.
Please give inverted ratings: 5 = strongly disagree, 4 = disagree, 3 = neutral, 2 = agree, 1 = strongly agree, 0 = Not

See research-based guidelines for usable design:
http://usability.gov/guidelines/guidelines_book.pdf

Chapter 6

Applicable)

The layout under each section of the website was consistent

See research-based guidelines for usable design:
http://usability.gov/guidelines/guidelines_book.pdf

Chapter 6

Perceived Visual Guidance

The colors helped me to distinguish different sections of the website

See research-based guidelines for usable design:
http://usability.gov/guidelines/guidelines_book.pdf

Chapter 14

The highlighted areas of a page helped me locate the information I needed

See research-based guidelines for usable design:
http://usability.gov/guidelines/guidelines_book.pdf

Chapter 14

I got to know the content of a page by skimming the highlighted areas

See research-based guidelines for usable design:
http://usability.gov/guidelines/guidelines_book.pdf

Chapter 14

Note: For the items indicated with * (reversed rating), please follow the instructions below the item.

As shown in Table 15, for each item with a negative score (e.g., less than the neutral or 3 rating), specific design actions can be initiated to address the identified issue. Whereas the specific redesign activity and identification of the actual requirements for improvements will depend upon a number of factors specific to the web project and constraints at hand, general guidance is provided to drive the design improvement of each issue that emerged from DEEP. When designers see a low score associated with an item, such as “I could quickly get to know the structure of the website by skimming its home page,” they have the opportunity to investigate, systematically check, and reconceive specific features of the web application concerning Structure and Information Architecture, possibly following established usability guidelines, heuristics, or design principles.

4.4.6. Overall Value of the Study

The importance of this study lies in the creation and validation of a valuable questionnaire developed to gather users’ perceived usability feedback around key design aspects of web-based systems. The articulation of the perceived usability construct into five design dimensions enables designers to gather usability feedback that more directly informs – in nature – the redesign activity. Neither the questionnaire instrument per se nor its results produce redesign requirements, but they do identify

user-perceived breakdowns of specific design aspects. This feedback is valuable as it provides a precise input as to how to make redesign decisions.

From a general perspective, DEEP takes the pillars of user perception (i.e., the usability phenotype) and makes that perception analytical in terms of using design dimensions to describe perceived usability. This allows me to bridge the gap between the usability genotype and analytical composition of the design (i.e., usability genotype; see Figure 23). DEEP can still be considered a user-based evaluation instrument, but, with respect to traditional approaches (e.g., usability testing or other self-reporting instruments), it captures the specific elements of the design at issue and allows designers to bridge the Evaluation-Design Alignment Gap by providing concrete evidence as to where to focus improvements.

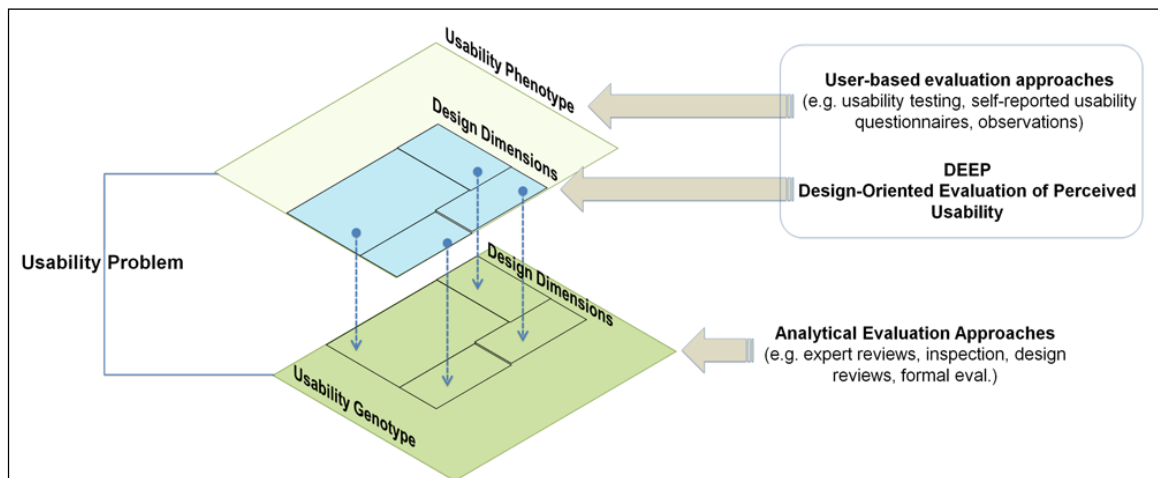


Figure 23. By mapping the perceived usability to key web design dimensions, DEEP addresses the Evaluation-Design Alignment Gap.

DEEP could serve as useful instrument for a broad range of HCI research studies. As DEEP provides a usability rating for items within specific design dimensions, this questionnaire will allow HCI researchers to conduct empirical studies that focus, for example, on the relationship between a design dimension (e.g., content, navigation, or graphics) and other factors of the user experience (e.g., trustworthiness, communicability, and joy). Questions such as “How does navigation usability affect trustworthiness?” and “How can content usability affect ‘flow’?” can be better answered

using DEEP to rate any analytical design dimensions (e.g., as independent variables) under investigation.

4.4.7. Limitations

Obvious limitations to this research exist, including the narrow scope of the design features and web application domain considered. I used university websites because they are a good example of large, information-intensive websites. The design of the chosen university websites was considered representative of the design dimensions under study. However, in order to strengthen the generalizability of the questionnaire, I would need to further validate DEEP to other relevant domains of information-intensive websites, including, for example, governmental portals, museums, other tourist attractions, and e-commerce sites.

Moreover, I recruited a controlled group of users made up exclusively of undergraduate students attending classes in one state in the Mid-western United States. Given this limitation, many surveyed users could possess other possible homogenous traits such as a relatively narrow age range, cultural perception, geographic similarity, and educational attainment. Further studies should be conducted regarding the DEEP effectiveness across user populations in order to gauge if or how usability of the tested design dimensions is perceived differently based upon varying user groups and for tasks of different natures.

This study focused on taking, as a starting point for selecting the measurement items, the most popular perceived usability questionnaires available (i.e., the most highly cited in the HCI literature). These questionnaires all come from reputable sources and have been extensively validated. Therefore, I feel strongly that the instruments utilized are representative of industry standards and best practices; however, given the potentially high and increasing number (and difficult accessibility) of ad hoc usability questionnaires developed in HCI research and professional arenas, I recognize that my starting basis of the measurement items might not be all-encompassing and is limited to the instruments reviewed in section 2.5.

Finally, although the in-depth, iterative analysis of the questionnaire's findings tested its reliability and validity, the long-term actual value of DEEP will result from having it used

extensively in its final version (Appendix K) on a number of websites. This will yield even further possible improvements and refinements to the instrument.

In next Chapter, the elements of the DEEP questionnaire and BREW evaluation approach were integrated in a controlled online experiment and a semi-structured interview to unpack the relationship between website usability and users' brand perceptions.

Chapter 5. Branded Interactions: Predicting Perceived Product Traits and User Image from Interface Consistency and Visual Guidance

Capitalizing on the methods used to evaluate online brand perception (Chapter 3) and design-based website usability (Chapter 4), this chapter investigates the relation between the two concepts (Yang & Bolchini, 2013). In particular, I explored whether and how specific design-based website usability attributes can influence how users perceive a specific brand attribute. This exploration was conducted through a rigorously-controlled online experiment and a follow-up semi-structured interview. The results revealed interesting differences between the usability attributes and gender groups.

5.1. Selecting the Most Reliable Design-based Usability Dimensions

The DEEP questionnaire captured five design-based usability dimensions: *content*, *information architecture*, *navigation*, *consistency*, and *visual guidance*. The correlations among consistency, information architecture, and navigation, however, were relatively high ($r > .66$). These correlations made it difficult to manipulate the usability of these dimensions separately on a website in order to compare their effects on brand perception. To tackle this issue, although my long-term goal was to see the effect of each dimension, as a starting point, this study only investigated the two most uncorrelated dimensions: *consistency* and *visual guidance*. The correlation between these dimensions was only .55. Exploratory factor analyses (Table 16) showed that the corresponding measurement items loaded strongly to two different factors, using either Varimax (assuming the two factors were orthogonal) or Oblique rotation (assuming the two factors were correlated). As such, consistency and visual guidance together can explain 77.93% of the total variation in perceived usability (with consistency 57.73% and visual guidance 20.20%, respectively).

Consistency is a key user interface principle broadly studied in usability engineering and interaction design (Fisher-Buttinger & Vallaster, 2006; Nielsen, 1994; Sharp, Rogers, & Preece, 2006). Visual guidance, although not explicitly defined, concerns the usable presentation of the elements in a user interface (Bolchini & Garzotto, 2008; Nielsen, 2006; Usability.gov, 2012). Following this tradition, below I have provided the operational definitions:

Table 16. Exploratory factor analyses of the measurement items for consistency (C) and visual guidance (V) (N=362) (Yang, Linder, & Bolchini, 2012).

Item	Varimax		Oblique	
	C	V	C	V
U1. I noticed abrupt changes in the layout of the pages.	.889		.947	
U2. The layout of the pages throughout the website was consistent.	.862		.860	
U3. The layout under each section of the website was consistent.	.832		.827	
U4. The highlighted areas of a page helped me locate the information I needed.		.891		.919
U5. I got to know the content of a page by skimming the highlighted areas.		.869		.903
U6. The colors helped me to distinguish different sections of the website.		.754		.755
Variance explained by each factor	57.73%	20.2%	57.73%	20.2%
Total variance explained	77.93%		77.93%	

Interface Consistency is a design property of a web application, by which the same user interface element (e.g., labels, icons, colors, or widgets) and their positions on the page represent the same navigational, symbolic, or functional behaviors throughout the system. Consistency is critical in regard to facilitating users' learnability and memorability of the system.

Visual guidance is a design property of a web application, by which texts, pictures, and widgets are grouped, colored, and positioned in order to maximize users' ability to recognize, read or locate relevant information.

From these definitions, it is clear that consistency is a cross-page attribute, whereas visual guidance is more of an intra-page concern.

5.2. Research Hypotheses

By manipulating the interface consistency and visual guidance of a content-intensive website, this study investigated how these two design-based usability factors could affect each of the attributes in Aaker's Brand Identity Planning Model (Chapter 2, section 2.1). Two hypotheses were tested.

H1: If two websites are significantly different on consistency, then the one with higher consistency yields a significantly more favorable perception of a brand identity attribute than the one with lower consistency.

H2: If two websites are significantly different on visual guidance, then the one with better visual guidance yields a significantly more favorable perception of a brand identity attribute than the one with worse visual guidance.

5.3. Methods

5.3.1. Choosing the Application Domain

The domain of this study is the online branding of US state universities. This domain was chosen for two reasons. First, I did not select commercial brands from the consumer products market because their brand identities are very different from each other. In contrast, universities share – with different emphases – many core and common brand attributes. For example, universities typically share the attributes of high-quality education, innovation within tradition, and commitment to student success (Wiese et al., 2009). By investigating university brands, the research outcomes can benefit multiple educational institutions rather than only helping one particular company. Second, commercial brands usually consist of a number of sub-brands, which may complicate the picture. For example, Proctor & Gamble owns about 80 sub-brands. The image of these sub-brands may influence each other.

5.3.2. Gauging Brand Perception

Aaker's (1996) original model was slightly adapted to create a survey instrument appropriate to use to gauge the brand perception of university websites. First, whereas the goods of a company and the company itself may represent different aspects of a brand, the brand as product and brand as organization constructs seemed inseparable for universities. In fact, the valuable “goods” of a university – the educational services it offers, such as teaching facilities, faculties, and research – and the university itself are highly intertwined. Therefore, I merged these two constructs and used brand as product to represent both a university itself and its educational services. Second, I decided to factor out the construct of brand as symbol because a symbol is more related to the visual property of the logo rather than to the usability of the website.

Then, based on the adjusted Aaker model, each of the brand identity sub-constructs were operationalized into more concrete attributes. For example, the *user image* of a university could be described by its students' personalities and social skills (Table 17). These brand identity attributes were collected, expanded upon, and consolidated from

existing literature on university branding (Aaker, 1996; Wiese et al., 2009) and through a focus group with 10 graduate students. In addition, the attributes were pre-tested through Amazon Mechanical Turk by asking participants (N=200) to rate their importance on a 7-point Likert scale from very unimportant to very important. The participants were asked: "If you are going to apply for a university in the United States, how important are the following factors for you?" Only the highly important attributes were kept. Through factor analyses, I found that brand as product consisted of three main factors: teaching and research quality, reputation, and student support. User image consisted of two factors: the personality and social skills of an excellent university student. Brand as person attributes clustered into one single factor. In the actual experiment, 7-point Likert scales from strongly disagree to strongly agree were used to capture brand as product and user image, while a variety of 7-point semantic differential scales were used to measure brand as person.

Table 17. Operationalize university brand identity into more concrete attributes.

Sub-Construct of Brand Identity	Brand Identity Attributes
Brand As Product (Measured by 7-point Likert scale; strongly disagree – strongly agree)	
• Teaching and Research Quality	B1. The university provides <i>high-quality teaching</i> B2. The university is committed to <i>academic excellence</i> B3. The university has <i>excellent teaching facilities</i>
• Reputation	B4. The university has a <i>unique identity</i> B5. The university has a <i>long history</i> B6. The university has a <i>high ranking in US News</i>
• Student Support	B7. The university <i>helps students to plan for their future</i> B8. The campus of the university is a <i>safe place</i> B9. The university offers a <i>personalizable path of study</i>
User Image (Measured by 7-point Likert scale; strongly disagree – strongly agree)	
• Personality	B10. They are <i>self-disciplined</i> B11. They are <i>hard-working</i> B12. They are <i>intelligent</i>
• Social Skills	B13. They have <i>excellent leadership</i> B14. They are <i>highly involved in social activities</i> B15. They are <i>the type of person whom I want to work with</i>
Brand As Person (Measured by 7-point semantic differential scales)	
	B16. Deceitful – Sincere B17. Boring – Exciting B18. Incompetent – Competent B19. Low-class – Upper-class B20. Charmless – Charming B21. Passive – Active B22. Weak – Strong

5.3.3. Experiment Design and Stimuli³

The experiment that I conducted was a 2×2 design with each independent variable varying on two levels: consistency (high/low) and visual guidance (good/poor) (Figure 24.a). The experiment was between-subject, so that each experiment group was exposed to only one of the four website versions (Appendix N):

- C+V+ (high consistency and good visual guidance);
- C-V+ (low consistency and good visual guidance);
- C+V- (high consistency and poor visual guidance);
- C-V- (low consistency and poor visual guidance).

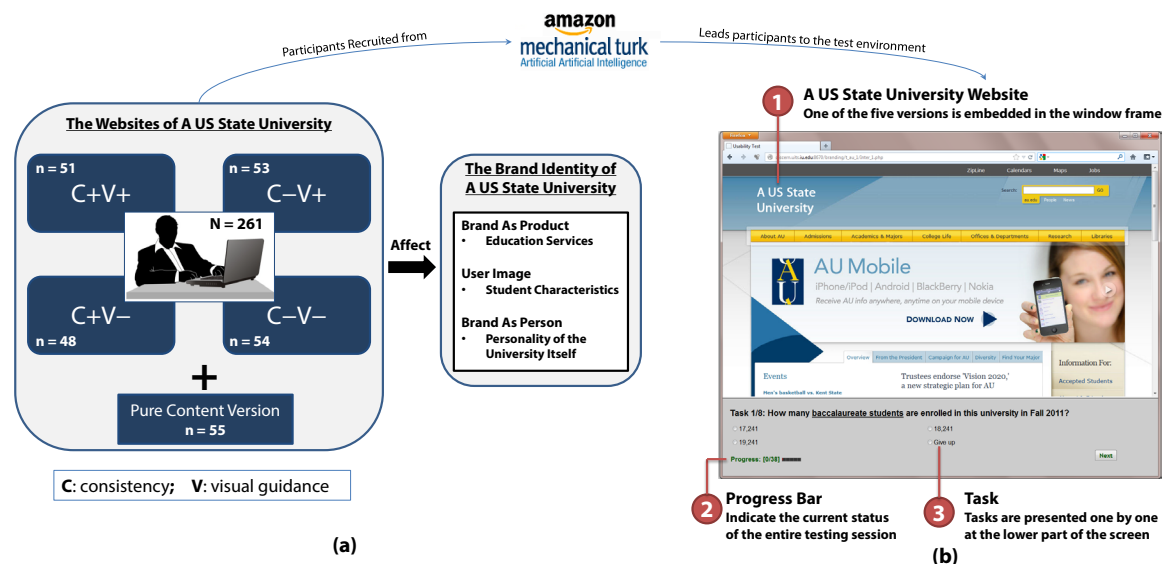


Figure 24. The experiment websites were embedded into the OFTEN environment and tested through Amazon Mechanical Turk.

In addition to these four main groups, I added a fifth group who was asked to use an online document that contained only the content of the experiment website. The aim was to examine whether the content itself was high quality, so that it would not negatively influence the results.

³ Related experiment materials can be found at: <http://discern.uits.iu.edu:8670/branding/>

In order to create the experiment websites, I began by pre-testing the usability of six candidate university websites using the DEEP questionnaire. The website that received the highest usability ratings was selected as the model for the experiment website.

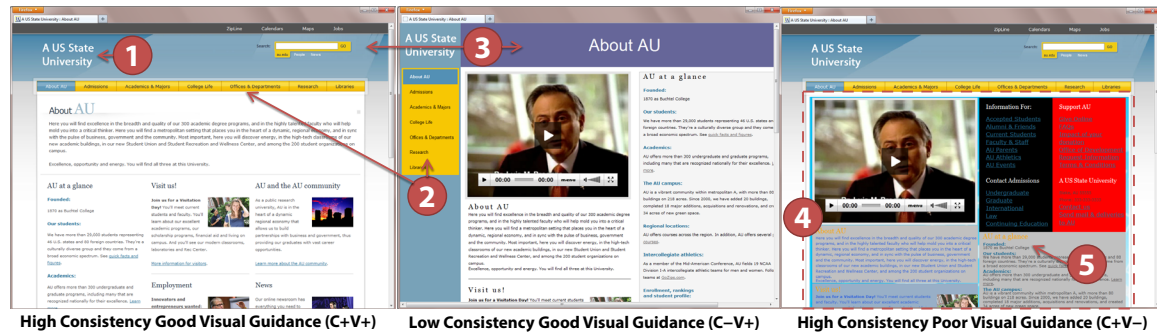


Figure 25. Examples of the experiment stimuli: ① a fictitious university name was used to control for prior brand knowledge; ② in the C-V+ version, the navigation bar was modified from horizontal to vertical to create inconsistent navigation; ③ in the C-V+ version, the color of the banner was changed to create inconsistent identity; ④ in the C+V- version, the pages were cluttered with poor background contrast and small fonts; ⑤ important content was placed where advertisements usually appear.

Then, in a period of six months, I modified, re-designed, and pre-tested the selected website to create the desired versions and achieve the intended perceived usability effects (Figure 25). The details of these research activities are described below.

C+V+ version: This version was very similar to the selected website with two main differences. 1) The university's identity was effaced by replacing the university's name with a generic name – a US State University – in order to avoid any bias caused by users' prior brand knowledge. Using fictitious brands was in line with the methodological approach of brand perception studies in electronic marketing (Dou et al., 2010; Lee, Rodgers, & Kim, 2009; Park & Lee, 2009; Wu, 2005). The related addresses and telephone numbers of the original website were replaced with fictitious content. 2) Several links were visible but disabled in order to prevent the users from navigating too far away from the main website sections.

C-V+ version: This version was created by worsening the consistency of the C+V+ version following existing usability guidelines. Two main consistency problems were created. 1) Navigation inconsistency (Nielsen, 2006). On a number of web pages, the

locations of or labels on the navigation panel were made different from those found on the homepage or higher level pages. 2) Identity inconsistency (Bolchini & Garzotto, 2008). When navigating through the website, users were expected to notice abrupt changes in color schemes and page alignment. Inside each page, however, the elements were well-organized and properly highlighted so that good visual guidance was kept. This version went through five rounds of revisions and pre-tests (20 participants in each round). In order not to make too brutal of changes on the first try, I started by creating moderate consistency problems only in the third and fourth level of the website architecture, while still keeping the interfaces of the first two levels consistent. Then, the usability of the modified site was pre-tested through Amazon Mechanical Turk. For the first three rounds of the pre-tests, however, no matter how much I increased the severity of the consistency problems, the users could not perceive the changes. Therefore, in the fourth round, the problems were also created in the second level of the site, which proved to be quite effective as perceived consistency dropped significantly. The implication for this change might be that users became less sensitive to inconsistent layouts as they got into the deeper levels of the site. The fifth pre-test was used to double check the reliability of the manipulation.

C+V- version: version: In this version, only the visual guidance was worsened. I created three main problems. 1) Highly cluttered pages (Bolchini & Garzotto, 2008; Nielsen, 2006). By imitating existing examples (Bolchini & Garzotto, 2008), I squeezed a lot of information into the web pages, using smaller fonts and poor color contrasts. 2) Important element with an insignificant location (Nielsen, 2006). On some pages, useful elements were put in the right column, where advertisements usually dwell. 3) Ad-like scrolling stoppers (Nielsen, 2006). On some of the long pages, I put a banner-ad at the bottom of the first screen, which pushed additional content “below the fold” and might prevent users from scrolling a page down (Nielsen, 2006). This version took three rounds of revisions and pre-tests (20 participants in each round) before the final version was accepted. In the beginning of the process, the changes were only made to the pages which were relevant to the user tasks (Table 18). This strategy, however, could also affect the consistency of the site whenever a user navigated to the *unchanged* sections. Therefore, in order to keep the website’s consistency level, the same design changes were applied to every page throughout the site. Although this decision could be

an over-manipulation – real-world problems may not be so pervasive on a site – it was a necessary compromise to the rigidity of a controlled experiment.

C-V- version: The changes made in C-V+ and C+V- versions were applied to this version. Therefore, it was poor on both aspects.

Pure content version: This version looked like a document. It contained the main content of the website, including texts and images, which were formatted like a survey report. It only had previous and next buttons and a simple dropdown menu for users to navigate through the content.

5.3.4. Procedure of the Online Experiment

The online experiment was conducted by combining the Amazon Mechanical Turk with the OFTEN field testing environment. Amazon Mechanical Turk was mainly used as a tool for recruiting participants. It indicated to the participants the URL used to access the experiment, which then took place in OFTEN. OFTEN was programmed to alternatively present the four website versions and the content-only version based on user ID, which guaranteed a controlled between-subject design.

Table 18. User tasks used in pre-tests and the actual experiment.

NO.	Task
1	How many baccalaureate students are enrolled in this university in Fall 2011?
2	Suppose you just graduated from high school and want to go immediately to college, please find the procedure for applying for an undergraduate program, and then, select from below the 3rd Step in the procedure.
3	To enrich students' college life, how many student organizations are in this University?
4	Suppose you are interested in applying for a bachelor degree in Biochemistry, please find out which college offers this major.
5	For students who want to live on campus, what is the lowest housing rate offered by this university?
6	Please check the research news and see who has received a big research grant on managing pipeline corrosion in Mexico.
7	Suppose you are interested in applying for a master degree in accounting, when is the application deadline for Fall enrollment?
8	You are planning to visit the campus of this university. Therefore, you want to call the Parking Services Office to ask for some parking tips. What is the phone number of the Parking Services Office?

Guided by OFTEN, each participant performed eight fact-finding tasks on either one of the four websites or using the pure content document. The tasks were designed based on the general services that a university website offers to occasional visitors (Bernier et al., 2002). For example, there were tasks that centered on finding admission information and faculty profiles (Table 18). Upon finishing the tasks, the participants who used the

website versions were asked to fill out both the perceived usability and brand perception questionnaires. The pure content group only filled out the brand perception questionnaires. Then, demographic information was collected and the participants received a reward code that they could submit back to Amazon Mechanical Turk to get their monetary incentive (\$2 for approximately 20 minutes of participation).

As to the demographic information, I followed suggestions from HCI user study methods (Sharp, Rogers, & Preece, 2006) and marketing survey techniques (McDaniel & Gates, 2010) to capture users' typical demographics traits including their gender, age, and education level⁴. Gender and age are two of the most common traits for segmenting consumer market (Solomon, 2008). Education level was involved because it is closely related to the experiment domain. It is possible that undergraduate and graduate students may see a university differently (e.g., the former may care more about education quality, whereas the latter may focus more on research resources). In addition, I also asked the participants' willingness to pursue a higher-level degree to see whether different motivation levels (i.e., prospective students may be more motivated than other users to have more serious considerations over the university) may yield different experiment results.

5.3.5. The Follow-up Semi-structured Interview

After analyzing the quantitative results from the online experiment, a follow-up semi-structured interview was conducted in order to further interpret the experiment findings. In particular, the main objective of this semi-structured interview was twofold. First, it was to double-check whether users could indeed perceive the interface consistency or visual guidance problems in a particular website intervention; and, second, to demystify *why* these problems were or were not able to affect users' opinions toward a specific aspect of the university brand.

The interview consisted of four user groups. Each group was exposed to one of the four website interventions (e.g., C+V+, C+V-, C-V+, and C-V-). A total of 20 student participants (12 females and 8 males, age > 18, 13 undergraduate and 7 graduate students) were recruited from Indiana University (IUPUI) through email invitations and a

⁴ The demographics questionnaire can be accessed through here:
<http://discern.uits.iu.edu:8670/branding/personalinfo.html>

university-level newsletter. Five participants were assigned to each user group. Since pervasive gender differences were found during the online experiment (see section 5.4), users' genders within each group were balanced: three females and two males were assigned to each group. More female participants were involved because the experiment manipulation turned out to be only effective for females.

The interview sessions took place in a typical usability lab at the Indiana University School of Informatics at IUPUI. In the lab, two desktop computers were set side-by-side on a long table. Upon arrival, the participants were asked to sit in front of one of the desktops. After introducing the participant to the general objectives of the study and guiding him or her to sign the informed consent, he or she was asked to perform eight fact-finding tasks on one of the experiment websites. This task-performance part was guided by the OFTEN environment, which was the same environment as used in the online experiment. Upon finishing the tasks, the interviewer would join the participant to help him or her fill in the brand perception questionnaire. Different from the online experiment, the participants were asked to take a break after finishing each section of the questionnaire (three sections: brand as product, user image, and brand as person). During each break, the interviewer reviewed each participant's rating on the items in that particular section and asked them the following types of questions:

You gave a high/low rating on [item name], why?

Do you remember which parts or elements on this website that made you feel like this?

Could you please point them out to me?

When asking these questions, the interviewer showed the website on the other desktop and gave the participant the mouse so that he or she could point to the website elements that he or she wanted to talk about. Users' activities on the screen and their voices were recorded using Morae Recorder (Version 3.1.1). Including the task-performance part, the entire session, on average, lasted around 40-45 minutes. Each participant was given a \$10 gift card as a compensation for their time and effort.

5.4. Results

Initially, 291 responses were received. Then, I polished them by removing the entries that left too many questions unanswered, those that showed apparent evidence of

gaming the testing system, and the extreme outliers. In the end, a total of 261 responses were retained: 51 belonging to the C+V+ group, 53 to the C-V+ group, 48 to the C+V- group, 54 to the C-V- group, and 55 to the pure content group. The respondents were balanced in gender (male = 143; female = 118). Most of their ages were within the 18 – 40 range (N = 219), while the rest were above 40 years old (N = 42). Most of the respondents were pursuing or had received an undergraduate degree (N = 201), while the rest either had not yet attended college (N = 25) or went to graduate school (N = 35). In addition, most of the respondents (N = 181) expressed their intention to pursue a higher level education.

5.4.1. Manipulation Checks

Manipulation checks were conducted by comparing the self-reported usability experiences (Figure 26) and task performances.

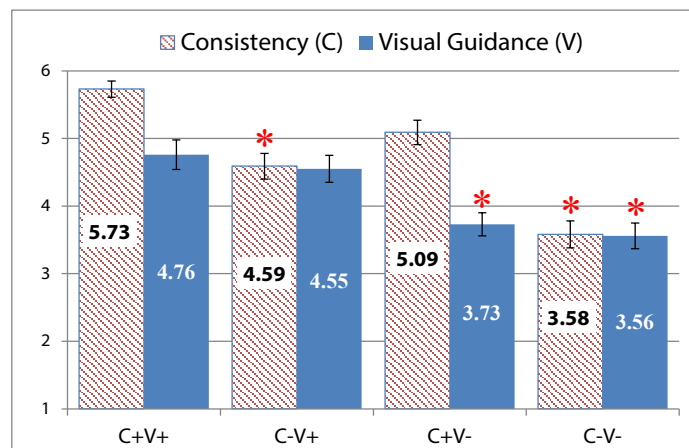


Figure 26. Successful manipulation checks of the website interventions. “*” indicates values that were significantly different from the C+V+ version ($p < .01$).

The scores for perceived consistency and visual guidance were calculated by averaging the corresponding usability indicators shown in Table 16. It was found that the perceived consistency scores were negatively skewed. Therefore, a logarithm transformation was applied to correct the skewness.

A one-way ANOVA (*Welch's* $F(3, 110.44) = 32.06, p < .01, \text{est. } \omega^2 = .31$) comparing the perceived consistency of the four websites suggested when compared to C+V+ ($M = 5.73, SE = .12$), perceived consistency decreased significantly in both C-V+ ($M = 4.59$,

SE = .19) and C-V- (M = 3.58, SE = .20), but not in the C+V- version (M = 5.09, SE = .18), indicating a successful experimental manipulation on consistency. A one-way MANOVA (*Wilks' λ* = .70, *F*(9,487) = 8.50, *p* < .01, η_p^2 = .11) was also conducted to compare the four websites on each of the three consistency indicators. It was found that all three indicators decreased significantly in the low consistency versions (C-V+ and C-V-), but not in C+V-, which was congruent with the overall findings. No interaction effect was found with the users' demographics traits.

The overall manipulations on visual guidance were also successful (ANOVA: *F*(3, 202) = 9.10, *p* < .01, ω^2 = .11). When compared to C+V+ (M = 4.76, SE = .22), perceived visual guidance dropped significantly in C+V- (M = 3.73, SE = .17) and C-V- (M = 3.56, SE = .19), but not in the C-V+ version (M = 4.55, SE = .20). When looking into each visual guidance indicator, it was found that the significant decrease occurred only in U4 (i.e., "The highlighted areas of a page helped me locate the information I needed," *p* < .004). The other two indicators showed the same trend (i.e., decreased substantially in the poor visual guidance versions), but not as strong.

As for users' performances, since the tasks that I asked users to accomplish were not difficult (i.e., basic fact-finding tasks), they were able to finish most of the tasks in spite of the less than satisfactory usability experiences in the worsened versions. The average task success rates were: C+V+ (M = 88%, SE = .02), C-V+ (M = 85%, SE = .02), C+V- (M = 85%, SE = .02), and C-V- (M = 83%, SE = .02). It might have happened by chance, but C+V+ version still yielded the highest success rate. Time-on-task showed a similar trend. Users spent significantly more time using C-V+ (M = 118 sec., SE = 7.3) than C+V+ (M = 91 sec., SE = 5.5) (*p* < .05). The time-on-task of C+V- (M = 105 sec., SE = 5.7), and C-V- (M = 112 sec., SE = 6.2) were also higher than that of C+V+, but the differences were not statistically significant. These results suggested that the usability problems either in consistency or visual guidance were not so severe that they greatly influenced users' performances in fact-finding. No significant interaction was found with users' demographics traits.

Based on the experimental manipulations, I coded consistency and visual guidance into two binary variables: consistency (0 = low, 1 = high); and visual guidance (0 = poor, 1 =

good) (Table 19). In this way, I could examine the main effects and interaction of these two usability factors on brand perceptions.

Table 19. Code consistency and visual guidance in two binary variables.

	C+V+	C-V+	C+V-	C-V-
Consistency	1	0	1	0
Visual Guidance	1	1	0	0

5.4.2. Brand as Product

All brand identity attributes can be found in Table 17. For the convenience of referencing, I also listed them below under the corresponding categories.

5.4.2.1. Teaching and Research Quality

- B1. The university provides *high-quality teaching*.
- B2. The university is committed to *academic excellence*.
- B3. The university has excellent teaching facilities.

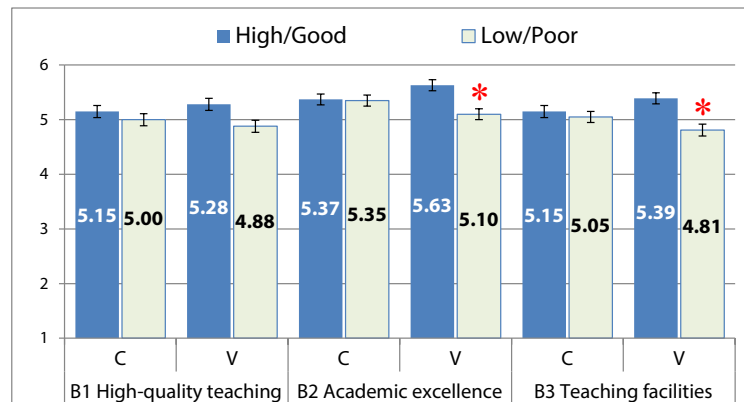


Figure 27. Worsened visual guidance (V) caused significant drop in B2 (Academic excellence) and B3 (Teaching facilities); but worsened interface consistency (C) did not seem to affect brand perception.

A 2×2 MANOVA (consistency \times visual guidance) suggested that the main effect of visual guidance on these three attributes was significant ($Wilks' \lambda = .92$, $F(3,200) = 5.74$, $p < .05$, $\eta_p^2 = .08$, Figure 27). The follow-up ANOVAs evaluating the effect of visual guidance on each attribute were significant at an adjusted significance level ($\alpha = .05/3 = .017$). Pairwise comparisons ($\alpha = .05/3/2 = .008$) revealed that the poor visual guidance websites yielded significantly lower perceptions on B2 Academic Excellence

($M = 5.10$, $SE = .10$) and B3 Teaching Facilities ($M = 4.81$, $SE = .11$) than the good visual guidance versions (B2: $M = 5.63$, $SE = .10$; B3: $M = 5.39$, $SE = .10$). In contrast, the effect of consistency and the interaction of consistency by visual guidance (consistency * visual guidance) were not significant.

A $2 \times 2 \times 2$ MANOVA with consistency, visual guidance, and gender (male, female) as the independent variables revealed a more precise pattern. Except that the effect of visual guidance remained significant, the interaction of visual guidance by gender (visual guidance * gender) was also statistically significant ($Wilks' \lambda = .92$, $F(3,196) = 5.68$, $p < .05$, $\eta_p^2 = .08$). As shown in Figure 28, females gave significantly lower ratings to all three brand identity attributes when visual guidance was worsened, but the males' ratings were almost the same between good and poor visual guidance conditions. In addition to gender, I also explored the other demographic traits, but no significant interaction was found.

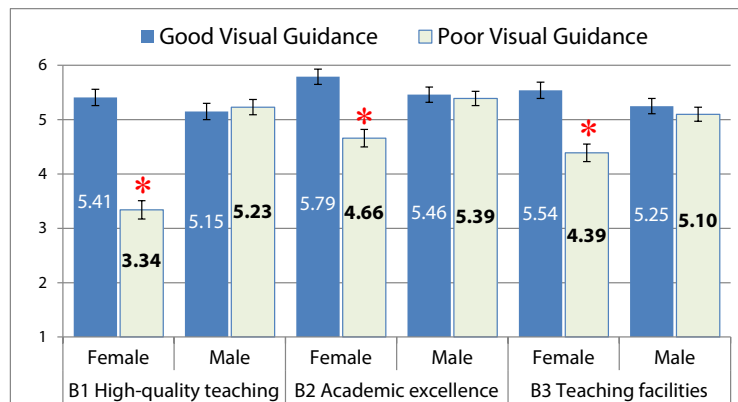


Figure 28. Females gave significantly lower ratings to all three teaching and research quality attributes when visual guidance was worsened, but the males did not (* $p < .05$).

5.4.2.2. Reputation

- B4. The university has a *unique identity*.
- B5. The university has a *long history*.
- B6. The university has a *high ranking in US News*.

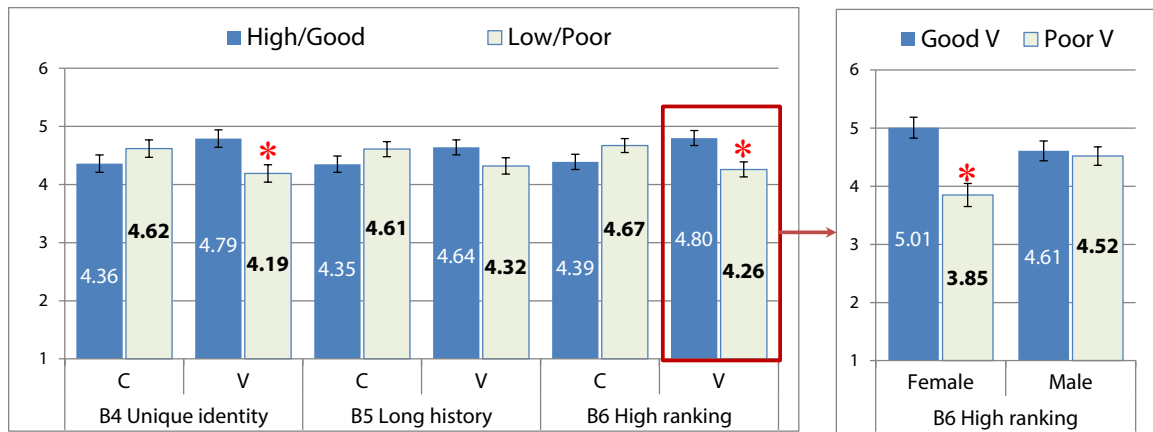


Figure 29. Worsened visual guidance caused significant drop in B4 (Unique identity) and B6 (High ranking, *female only*); but worsened consistency did not seem to affect brand perception.

A 2×2 MANOVA testing the main effect of visual guidance on these three attributes was significant ($Wilks' \lambda = .94$, $F(3,200) = 4.18$, $p < .05$, $\eta_p^2 = .06$). The follow-up ANOVAs and pairwise comparisons suggested that B4 Unique Identity ($M = 4.19$, $SE = .15$) and B6 High Ranking ($M = 4.26$, $SE = .13$) dropped significantly when visual guidance was worsened, when compared to the good visual guidance versions (B4: $M = 4.79$, $SE = .15$; B6: $M = 4.80$, $SE = .13$). The main effect of consistency and the interaction of consistency * visual guidance were not significant.

The interaction of visual guidance * gender was significant on B6 High Ranking ($F(1,198) = 9.12$, $p < .017$, $\eta_p^2 = .04$). Females gave it significantly lower ratings in the poor visual guidance conditions ($M = 3.85$, $SE = .20$) than in the good visual guidance conditions ($M = 5.00$, $SE = .16$). The males' ratings did not show much difference (Figure 29).

5.4.2.3. Student Support

- B7. The university helps students to plan for their future.
- B8. The campus of the university is a *safe place*.
- B9. The university offers a personalized path of study.

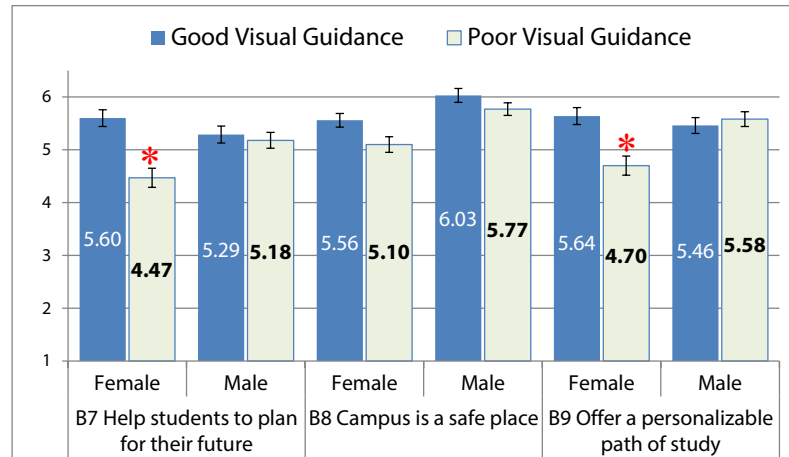


Figure 30. Females gave significantly lower ratings to B7 (Help students to plan for their future) and B9 (Offer a personalized path of study) when visual guidance was worsened, but the males did not (* $p < .05$).

Similarly, only visual guidance showed a significant main effect on these attributes ($Wilks' \lambda = .94$, $F(3,200) = 4.08$, $p < .05$, $\eta_p^2 = .06$). Good visual guidance yielded significantly higher ratings on B7 Help Students to Plan for Their Future ($M = 5.44$, $SE = .11$), than did the poor visual guidance websites ($M = 4.89$, $SE = .12$). The interaction of visual guidance * gender was significant on B7 Help Students to Plan for Their Future ($F(1,198) = 12.84$, $p < .017$, $\eta_p^2 = .05$) and B9 Offer a Personalized Path of Study ($F(1,198) = 14.01$, $p < .017$, $\eta_p^2 = .05$). Again, only females were significantly influenced by visual guidance in perceiving these brand identity attributes (Figure 30).

5.4.3. User Image

Under this dimension, the participants were asked to rate the *characteristics of the university's students*, namely personality and social skills.

5.4.3.1. Personality

- B10. Students of this university are *self-disciplined*.
- B11. Students of this university are *hard-working*.
- B12. Students of this university are *intelligent*.

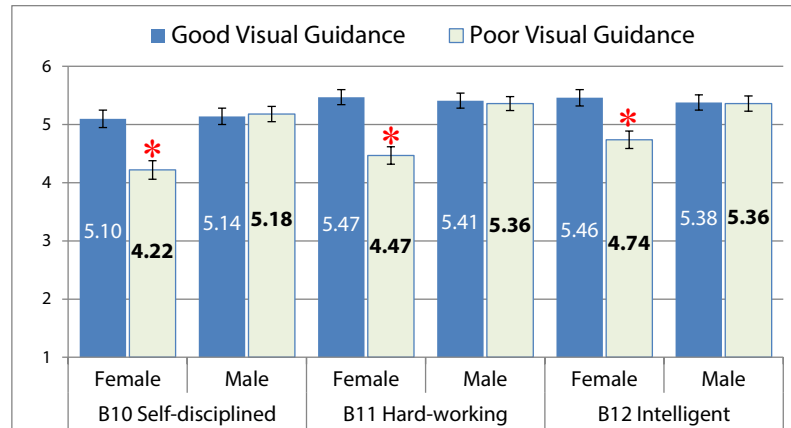


Figure 31. Females gave significantly lower ratings to all three attributes about the personality of typical students when visual guidance was worsened, but the males did not (* $p < .05$).

Visual guidance had a significant main effect on B11 Hard-Working ($F(1,202) = 10.58$, $p < .017$, $\eta_p^2 = .05$). It received significantly lower ratings from the poor visual guidance websites ($M = 5.00$, $SE = .10$) than from the good visual guidance websites ($M = 5.45$, $SE = .10$, $p < .008$). The main effect of consistency was not significant. The interaction of visual guidance * gender ($Wilks' \lambda = .93$, $F(3,196) = 4.61$, $p < .05$, $\eta_p^2 = .07$) suggested that, with poor visual guidance, the perception of all three attributes dropped significantly for females, but not for males (Figure 31).

5.4.3.2. Social Skills

- B13. The students have *excellent leadership*.
- B14. They highly involved in social activities.
- B15. They are the type of person I want to work with.

Both the main and the interaction effects of consistency and visual guidance on these attributes were not significant ($p > .05$). This time, however, gender difference completely suppressed the impact of visual guidance. In fact, for the females, all three attributes decreased significantly in the poor visual guidance conditions ($p < .05$) (Figure 32). In contrast, it was quite surprising to find that the males' ratings did not even show a decreasing trend.

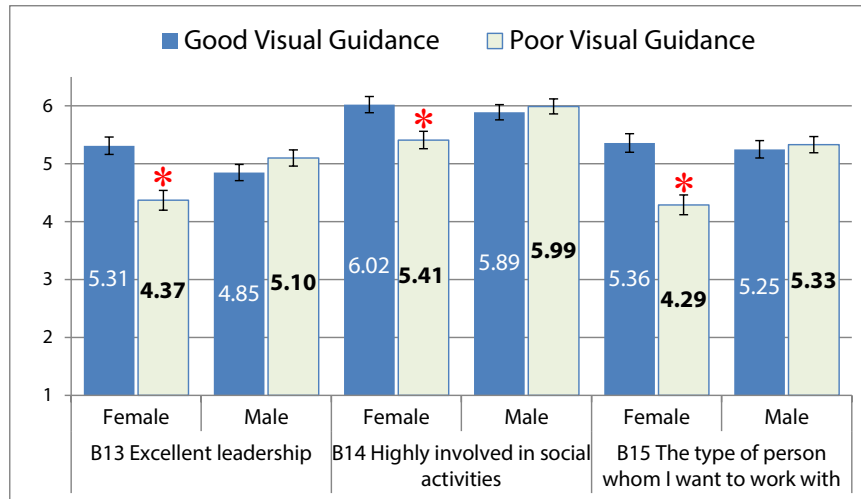


Figure 32. In the presence of poor visual guidance, female users gave significantly lower ratings to all three attributes about the social skills of typical students (* $p < .05$).

5.4.4. Brand as Person

Under this dimension, the participants were asked to consider the university as *a person* and express their opinions toward him or her.

- B16. Deceitful – Sincere
- B17. Boring – Exciting
- B18. Incompetent – Competent
- B19. Low-class – Upper-class
- B20. Charmless – Charming
- B21. Passive – Active
- B22. Weak – Strong

Although the MANOVA test indicated significant effects of both consistency and visual guidance on these attributes, all of the follow-up tests were non-significant. By controlling for the main effect of gender, I found that the ratings on B17 (Boring – Exciting) were significantly lower in the poor visual guidance conditions ($M = 3.86$, $SE = .16$) than in the good visual guidance conditions ($M = 4.52$, $SE = .15$).

5.4.5. Investigating the Gender Difference on Brand Perception

The analyses above indicated that visual guidance has a significant effect on brand perception that was highly gender-specific. In general, only females were influenced by

the worsened visual guidance when evaluating the presented brand. In order to investigate the potential reason for this gender difference, the following analyses were conducted.

First, a conceptual path model illustrating the experiment was established (Figure 33). Based on the experiment design, the manipulations on website visual guidance were supposed to be perceived by the users first, which, in turn, would affect their brand perceptions – a full mediator effect. However, in addition to this mediator effect, there was still a possibility that the experiment manipulations would have additional direct effects on brand perceptions. The reasons for these direct effects could be that the manipulations involved additional changes besides visual guidance that could affect brand perceptions.

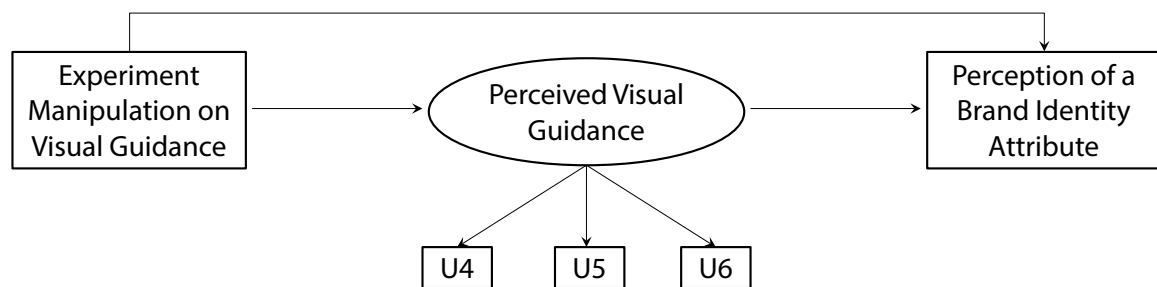


Figure 33. A conceptual path model of the experiment.

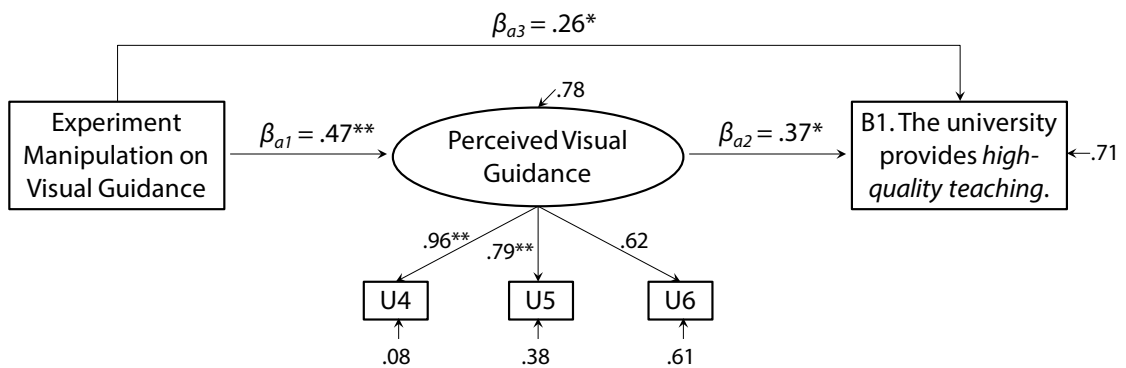
Accordingly, for each significant visual guidance * gender interaction detected in the previous sections ($N = 12$), the path coefficients in the model were estimated using structural equation modeling technique for females and males separately to see what had caused the gender differences. A *critical ratio for differences test* was used to compare the path coefficients in the female and male models. The Amos 5 software was used to conduct these tests.

Taking B1 (High-quality teaching) as an example, a two-group path model was established (Figure 34). The model showed that the manipulation on visual guidance significantly predicted perceived visual guidance (see β_{a1} and β_{b1}), and there was no significant difference between females and males (based on the critical ratio for differences test, $p > .05$). Consistent with the manipulation checks, this again suggested that females and males perceived the experiment manipulations equally. Therefore, the

cause of the gender difference was unlikely to be that the males failed to perceive the changes in visual guidance.

Next, perceived visual guidance significantly predicted the perception on B1 (see β_{a2} and β_{b2}) and no significant difference was found between females and males ($p > .05$). Thus, the females' and males' opinions toward B1 seemed to be influenced equally by their perceptions of the visual guidance. In other words, it was also unlikely that the males did not take into consideration website visual guidance when they assessed B1. Up to now, the reason of the gender difference was still unknown.

Female:



Male:

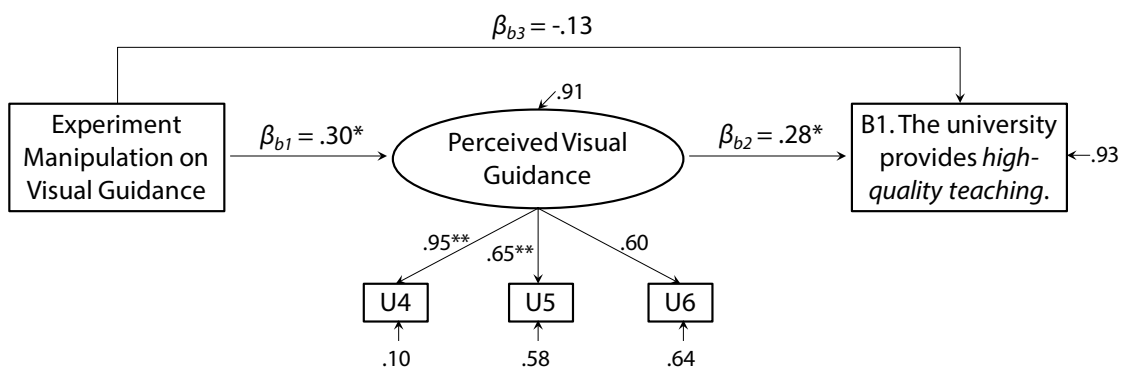


Figure 34. Two-group path analysis on B1 (High-Quality Teaching); significant difference was found between β_{a3} and β_{b3} . ($*p < .05$, $**p < .001$).

In contrast, when comparing the third path (manipulation → B1), it was found that the manipulation on visual guidance could significantly predict the females' perceptions on B1, but not the males' perceptions. In addition, the critical ratio for differences test

showed that the path coefficient in the females' model ($\beta_{a3} = .26$) was significantly higher than that in the males' model ($\beta_{b3} = -.13$, $p < .05$). This finding indicated that there were *additional attributes* that were affected by the experiment manipulation, and these additional attributes could *only* influence the females' opinions toward B1.

This two-group path model showed a good fit to the data ($\chi^2(8) = 13.96$, $p = .08$). The Goodness-of-Fit Index (GFI = .97), Adjusted Goodness-of-Fit Index (AGFI = .90), Normal Fit Index (NFI = .95), and Comparative Fit Index (CFI = .98) were all higher than the .90 threshold. The Standardized Root Mean Square Residual (SRMR = .07) was lower than the .10 cut-off. The only deficiency was that the Root Mean Square Error of Approximation (RMSEA = .06), which was slightly higher than the .05 cut-off. Therefore, overall, this two-group model accurately reflected what had happened in the experiment. In addition to the path analysis, I also used Baron and Kenny's four-step criteria (Baron & Kenny, 1986; Frazier, Tix, & Barron, 2004) to examine the mediator effects by performing four regression analyses for each gender group: 1) the manipulation on visual guidance should significantly predict the perception on B1; 2) the manipulation on visual guidance should significantly predict perceived visual guidance; 3) perceived visual guidance should significantly predict the perception on B1; and 4) using both manipulation and perceived visual guidance to predict B1, the effect of manipulation on B1 should drop substantially compared to 1). For females, all four regression analyses yielded the expected outcomes listed above, indicating that the effect of the manipulation on B1 was indeed mediated by perceived visual guidance. For males, however, criterion 1) was not met: the manipulation itself did not significantly predict B1. Although Zhao et al. (2010) argued that criterion 1) might not be necessary for a significant mediator effect, it was still possible that the effect was spurious for males.

Besides B1, seven other brand identity attributes (B2, B3, B7, B9, B10, B11, and B15) suggested the same pattern in their corresponding path models (Table 20). In all of these models, manipulation had a significant influence on perceived visual guidance, while perceived visual guidance, in turn, significantly influenced brand perception. These two types of effects were equal for females and males. However, the direct effect of manipulation on brand perception was statistically significant only for females and significantly higher than that for males. The fit indices of these models can be found in the Appendix O. All of these models demonstrated satisfactory model fit.

Table 20. Analyzing the visual guidance * gender interaction (* $p < .05$, ** $p < .001$; CRD: Critical Ratio for Differences; '=': no significant difference; '>': significantly higher for females).

	Manipulation → Perceived VG			Perceived VG → Brand Attribute			Manipulation → Brand Attribute		
	Female (β_{a1})	Male (β_{b1})	CRD	Female (β_{a2})	Male (β_{b2})	CRD	Female (β_{a3})	Male (β_{b3})	CRD
B1	.47**	.30*	=	.37*	.28*	=	.26*	-.13	>
B2	.47**	.30*	=	.29*	.21*	=	.37**	-.03	>
B3	.47**	.29*	=	.40**	.30*	=	.29*	-.02	>
¹ B6	.47**	.30*	=	.12	.08	=	.38**	.00	>
B7	.46**	.30*	=	.35*	.33*	=	.30*	-.05	>
B9	.47**	.30*	=	.28*	.22*	=	.25*	-.12	>
B10	.47**	.30*	=	.33*	.37**	=	.23*	-.12	>
B11	.47**	.30*	=	.26*	.24*	=	.36**	-.05	>
¹ B12	.47**	.30*	=	.34*	.25*	=	.20	-.06	=
¹ B13	.46**	.30*	=	.45**	.31*	=	.21*	-.21*	>
¹ B14	.46**	.30*	=	.26*	.12	=	.16	-.09	=
B15	.47**	.30*	=	.40**	.40**	=	.24*	-.16	>

¹ These attributes have a different pattern from the others.

The rest of the four brand identity attributes showed certain variations to the major pattern discussed above (Table 20). For both genders, B6 (High Ranking) was not significantly predicted by perceived visual guidance ($\beta_{female} = .12$; $\beta_{male} = .08$). The only reason for the significance of the guidance * gender interaction was that the direct effect of the manipulation on B6 was statistically significant only for females ($\beta_{female} = .38$) and significantly higher than that for males ($\beta_{male} = .00$).

The model for B13 (Excellent leadership) had almost the same pattern as the major trend. The only difference was that the manipulation had a significant *negative* effect on B13 for males ($\beta_{male} = -.21$). This result suggested that the additional changes in the experiment manipulation had an opposite impact on females ($\beta_{female} = .21$) and males in perceiving B13.

B12 (Intelligent) and B14 (Highly involved in social activities) were difficult to interpret. They did not show significant differences in any of the three paths. The only explanation to the significance of the visual guidance * gender interaction in these two attributes could be that all three paths in the females' models had higher absolute values than those in the males' models. However, this interpretation should be taken with caution.

In spite of the slight variations, when the mediation effect through perceived visual guidance was statistically controlled, the analyses above showed quite a strong trend that the experiment manipulation had additional significant effects on brand perceptions for females. Although the mediation effect was, in general, stronger than the additional effects (by comparing the path coefficients in each model), it was the combination of these two types of effects that caused females to give significantly lower ratings to a number of brand identity attributes in the poor visual guidance conditions. In contrast, even though the mediation effects on males did not show much difference from that on females (it might be spurious), the males' brand perceptions were not significantly changed.

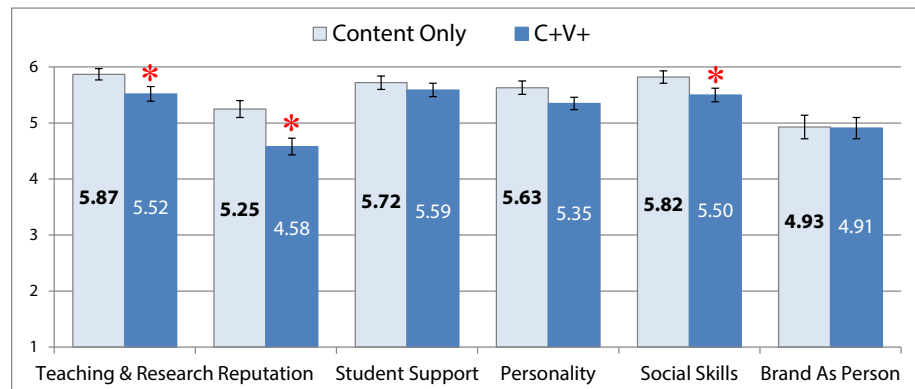


Figure 35. The pure content document yielded the most satisfactory perceptions on all brand identity factors (*significantly lower than the pure content group, $p < .05$).

5.4.6. Content Only and Full Website

Most of the brand identity attributes received the highest ratings from the group who used the pure content document (without any significant gender differences). Figure 35 demonstrates the comparisons of the average ratings on each brand identity factor between the pure content document and the C+V+ version. Independent samples t-tests suggested that the pure content document yielded significantly higher ratings on teaching and research quality, reputation, and social skills than C+V+. These results first indicated that the content that this university website conveyed was of high quality. Therefore, it could not negatively influence the research results. Second, although C+V+ was considered to be good on both consistency and visual guidance, it still reduced the branding effect of the content itself. Therefore, even the C+V+ version could still be further improved upon in terms of branding effectiveness.

5.4.7. Findings from the Follow-up Interviews

After transcribing the interviews, I analyzed the transcribed text in four aspects. 1) Did the users perceive the interface consistency or visual guidance problems? 2) What were the users' overall impressions toward the university? Were their opinions affected by the interface consistency or visual guidance problems? 3) Based on which aspect(s) of the website did the users give ratings to the brand identity attributes? Were their opinions affected by the interface consistency or visual guidance problems? 4) Which aspects did users think were the most important aspects of a university's website?

1) Did users perceive the interface consistency or visual guidance problems?

The users perceived the interface consistency and visual guidance problems while performing the eight tasks and accurately described them in the interviews. When I asked "Was there anything that you liked or disliked about this website," the users who used the low consistency versions described both the navigation inconsistency and identity inconsistency problems. The following are two quotes from the users:

"The design completely changed when you go... when you kind of drill down to specific area. That can be kind of distracting because you are looking at that one format and then it did not fix... The site just completely changed." **[Navigation Inconsistency]**

"The design styles are also not uniform. When I clicked in, I felt like it's a different university." **[Identity Inconsistency]**

Users of the poor visual guidance versions also accurately described the three types of visual guidance issues:

"I think this information is... is a lot of information in this small space. So it's kind of overload." **[Cluttered Page]**

"I used to see this (supporting links) somewhere at the bottom. Its location should have more pertinent things than how to support the university." **[Unimportant Content at Important location]**

"You know, a lot of the pages were scroll-down pages... I know that's an easier way to build a website, but people may miss information that way. Sometimes I forgot to scroll down and sometimes if you don't scroll down

far enough you may miss information that you are looking for.” **[Long Page and Scroll Stoppers]**

These results again showed that the experiment manipulations were successful.

2) What were the users’ overall impressions toward the university? Were their opinions affected by the interface consistency or visual guidance problems?

The users were asked what they thought of the university overall after using its website. Their answers are summarized in Table 21. Even with such a small sample, it is still possible to see that females were more inclined to give harsh comments about the university than males when the website had poor visual guidance. In particular, in the good visual guidance conditions, all of the users considered it a good or intermediate-level university, no matter whether the website had inconsistent interfaces. In contrast, in the poor visual guidance condition, three of the six female users thought that the university was a lower level university. None of the males, however, gave it such a low rating.

Table 21. Users’ overall impressions toward the university (numeric data stand for the number of participants who gave a particular rating).

		Good Visual Guidance		Poor Visual Guidance	
		Female	Male	Female	Male
High Consistency	Good	0	1	0	0
	Intermediate	3	1	2	2
	Low	0	0	1	0
Low Consistency	Good	1	1	0	0
	Intermediate	2	1	1	2
	Low	0	0	2	0

The major reason why the three females ranked the university as low was that the poor visual guidance yielded problems for them in regard to accessing information. For example, when asked whether they would consider applying to the university, one of the female users compared the website to her own university’s website and said the following:

“Honestly no (i.e., would not consider applying for the university). It's confusing... Coming to my school's website, it seems like, wow, it's good. It has a lot of information over there, whereas coming to this [website], I don't see much information and I couldn't find any information I needed... I'll be comparing, if I compare my school's website with this, I'll go to my school, not this.”

In fact, her university's website contains an approximate equal amount of information as the experimental website. The reason why she said that she "didn't see much information" was probably because the experimental website placed crucial information at unimportant locations and the scroll stoppers prevented her from scrolling down to read information below the fold:

"When I go to this place, usually I look for some information maybe on this side but usually I need to go down this way. I expect the page to change... but after answering two three questions, I notice that I'm getting answers over here not over here."

Similarly, another female user said that, in the real world, she might never apply for a university with such a website, because *"it had too much information which was not arranged"* and she thought that the *"best university should have the best website."* She also assumed that a good university should have a big group of people to maintain and update the website. The third female user mentioned that she was not able to accomplish a task because she did not think that the position where the answer was located could contain useful information. In addition, she thought that the colors and the styles of the interfaces were just like the websites of the lower level universities that she had run into before.

In contrast, the male users were reluctant to say that the university was not good, even though they encountered the same problems as the female users. For example, one of the male users said that he would not judge a university solely on its website.

"There is a lot of research [that] goes with that (i.e., whether the university is worth applying for). I would definitely say if two [universities] are really close like by rankings and everything and I don't know which one I like more, I think a better website could possibly influence [me]. But, I wouldn't pick one based solely on [its] website."

Another male user thought that the influence of a website was quite limited:

"You know, obviously it (e.g., the visual guidance problems) is not the most critical thing. Could be annoying, yeah, could be. It really has some effect... You hope it wouldn't be the most important thing, but, on the other hand, if you are looking at several universities and you have never heard of this one or it is somewhere in another state probably... If I

cannot find information I want, I just give up at some point of time... but I don't think that would be a determining factor.”

“The first thing I would be doing is to look and see if they offer the major I'm looking for. It'd be better to find that pretty quickly.”

In sum, the female users were more prone to relating the visual guidance problems of the website to the overall quality of the university, whereas the male users considered it more of a peripheral factor.

Interface consistency, in contrast, did not seem to either affect the female users' or male users' overall impressions toward the university. Most of the users did not mention the consistency problems when explaining why they thought the university was good, intermediate, or low level. When explicitly asked whether the inconsistent interfaces gave them negative feelings toward the university, one of the female users in the C-V+ group said: *“It may not be a problem with the university, but it seems like they have different development teams and they do not use consistent themes.”* The other two female users in the C-V+ group simply answered that they did not care whether the interfaces were consistent when judging the quality of a university. A male user in the C-V+ group thought that consistent layouts were something good to have, but not necessary: *“It is a luxury thing. If you have it, it is good. If you do not have, it is not really that much.”*

3) Based on which aspect(s) of the website did the users give ratings to the brand identity attributes? Were their opinions affected by the interface consistency or visual guidance problems?

The reasons that the users gave about their ratings on specific brand identity attributes can be clustered into three categories: content, visual design, and accessibility. The full list of reasons can be found in Appendix P. Typically, users did not provide a reason for each attribute. They, in general, elaborated upon their opinions when assessing the first several attributes, and then repeated the same reasons for the rest of the attributes.

Table 22 shows the number of reasons in each category for a particular website version. Website content (i.e., the text) was the primary resource used by users to judge or deduce how much a brand identity attribute was likely to be true.

Table 22. Number and types of reasons for particular ratings on specific brand identity attributes.

		Good Visual Guidance		Poor Visual Guidance	
		Female	Male	Female	Male
High Consistency	Content	19 (14, 1, 4)*	13 (12, 0, 1)	21 (16, 0, 5)	8 (5, 1, 2)
	Visual Design	12 (5, 3, 4)	9 (7, 0, 2)	5 (2, 0, 3)	9 (0, 3, 6)
	Accessibility	3 (3, 0, 0)	1 (1, 0, 0)	2 (1, 0, 1)	4 (1, 0, 3)
Low Consistency	Content	9 (8, 0, 1)	6 (5, 0, 1)	18 (8, 1, 9)	12 (9, 1, 2)
	Visual Design	1 (1, 0, 0)	4 (1, 0, 3)	9 (2, 0, 7)	4 (1, 0, 3)
	Accessibility	1 (0, 1, 0)	4 (4, 0, 0)	3 (0, 0, 3)	0 (N/A)

*Total number of comments (Positive comments, Neutral comments, Negative comments) → 19 (14, 1, 4)

Among all reasons that the users gave, 59.89% (106 out of 177) of the reasons referred to the website content. For example, a user gave a positive rating to the attribute *high-quality teaching* (B1), because he saw in the “quick facts” section of the website that “a large number of students are enrolled in the university.” He said: “not many students would go there if the university does not provide high-quality teaching.” Another user thought that the students of the university were not *the type of person she wants to work with* (B15), because she saw that the university had too many student organizations and extracurricular activities: “In regard to working, I like people with a good work ethic so that they can put more effort into the work” [not into extracurricular activities].

The second user in the previous paragraph, however, was more of an outlier. That is because most of the users gave negative ratings to brand identity attributes not because the content that they read gave them bad feelings, but because they were not able to find the relevant content. For example, a female user in the C–V– group thought that the university was unlikely to *help students to plan for their future* (B7) because she did not see anything about *career services*. This finding, however, does not mean that people who gave positive ratings saw the relevant content. Some of the brand identity attributes were not mentioned on the website (across all versions), but still received positive ratings. This phenomenon can be at least partly explained by the website’s visual guidance.

For the websites with visual guidance, 17.86% (5 out of 28) of female users’ content-related comments (reasons) were negative, which increased to 35.90% (14 out of 39) for

the websites with poor visual guidance (Table 22). The poor visual guidance might have made the users think that it was difficult to find information on the website in general. Therefore, they ascribed the fact that they did not see relevant content to the visual guidance problem of the website – the content might have been there, but the website made it difficult to find. In contrast, information was easy to find on the websites with good visual guidance. When asked to rate an attribute that they knew nothing about, the users did not think that it was the problem of the website that limited their knowledge, but because they just happened not to see the relevant content. Therefore, they either gave these attributes neutral ratings or tried to deduce a rating from the other content that they saw.

For the male users, 11.76% (2 out of 17) of their content-related reasons were negative in the good visual guidance conditions, which increased to 28.57% (4 out of 14) in the poor visual guidance conditions. Although this increase was relatively big in terms of percentage, only four negative reasons were gathered in the poor visual guidance conditions, which even less than the number of negative reasons given by the female users in the good visual guidance conditions. This finding might be another indication that the male users were less prone to relate visual guidance problems to the brand image of a university.

Next, 29.94% of the reasons why users gave certain ratings to the brand identity attributes (53 out of 177) referred to the *visual design* of the website. The number of negative reasons increased in the poor visual guidance conditions for both females (from 30.77% to 71.43% or from 4/13 to 10/14) and males (from 38.46% to 69.23% or from 5/13 to 9/13). These visual design issues, however, were not related to visual guidance, but to graphic design style of the website. For example, a female user in the C+V– group did not think that the university had a unique identity (B4) because “the design of the website was fairly generic.” A male user in the C–V– group thought that the university did not have a long history (B5) because “the website seemed fairly modern, fairly contemporarily thought-out.” Thus, the poor visual guidance erased the unique characteristics of the graphic design and made the website look like it had been recently created, which resulted in the inferior ratings on the two brand attributes.

There were only a few reasons that were related to website accessibility (10.17%, 18 out of 177). For example, one male user in the C+V- group stated that because he could not find the on-campus housing information, he gave low ratings to all of the attributes related to *student support* (B7, B8, and B9). However, because the cases were too few, it was difficult to use them to establish a trend. On the other hand, this finding suggested that users seldom directly relate the meaning of a brand identity attribute to website accessibility.

In contrast, interface consistency seemed to have certain effects on users' opinions, but not as strong as visual guidance. For females, the negative reasons that they gave regarding website content and visual design were 22.50% (9 out of 40) and 41.18% (7 out of 17), respectively, of the total number of reasons that they gave for the high consistency conditions. Although these percentages increased to 37.04% (10 out of 27) and 70.00% (7 out of 10) in the low consistency conditions, the absolute number of negative reasons did not increase. Similarly, for males, the negative reasons that they gave regarding website content and visual design were 14.29% (3 out of 21) and 44.44% (8 out of 18), respectively, of the total number of reasons that they gave in the high consistency conditions. These percentages increased to 16.67% (3 out of 18) and 75.00% (6 out of 8) in the low consistency conditions, but again, the absolute number of negative reasons did not change. These findings suggested that low interface consistency might make users provide less positive reasons, but would not encourage them to give more negative reasons.

4) Which aspects did users think were the most important aspects of a university's website?

When interpreting the mediation models created in section 5.4.5, I found that *website aesthetics* might be the factor that mediated the additional direct effect of visual guidance manipulation on brand perceptions (see 5.5.3). In order to collect preliminary evidence about this assumption, I asked the users to rank the importance of three website attributes (i.e., *website content*, *ease of finding information*, and *aesthetical design*) to the image of a university. On a scale of one to three (the smaller the rating, the more important it is), females gave almost equal importance to the three attributes: website content ($M = 1.67$, $SD = .33$), ease of finding information ($M = 1.75$, $SD = .16$),

and aesthetical design (M = 1.56, SD = .29). In contrast, males thought that *ease of finding information* was the most important factor (M = 1.29, SD = .18), *website content* the second most important factor (M = 1.75, SD = .25), while *aesthetical design* the least important factor (M = 2.71, SD = .18). If the direct effects in the models were indeed mediated by website aesthetics, these rankings might explain why the effects were not statistically significant in the males' models. As stated by one of the male users: "To me, it [aesthetics] is not really [that important]. As long as it [the website] has what I need and is easy to use, I kind of overlook that."

5.5. Discussion

5.5.1. No Effect of Consistency on Brand Image

Interface consistency, as mentioned earlier, is one of the most noted usability criteria in existing interaction design guidelines. In addition, low consistency was the only usability problem in this study that significantly worsened users' task performances. The inconsistent layouts disoriented the users and made them spend additional time trying to find the desired information (i.e., the average time-on-task increased by 29.67% in the C-V+ version). Surprisingly, however, it did not have a strong effect on the perception of any of the brand identity attributes. Table 23 shows the correlation of perceived consistency with all 22 brand identity attributes. The highest correlation was only .32, indicating that a significant decrease in consistency may only cause very slight drop in brand perception.

Table 23. Correlations between perceived consistency and each brand identity attribute (*p < .05; **p < .01).

B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11	B12	B13	B14	B15	B16	B17	B18	B19	B20	B21	B22
.32**	.25**	.30**	.19**	.05	.12	.24**	.10	.11	.18*	.12	.11	.23**	.16*	.17*	.21**	.23**	.27**	.12	.19**	.18**	.23**

These low correlations, the online experiment findings, and the follow-up interviews all suggest that interface consistency has minimal relation with the image of a brand in users' minds. However, instead of simply concluding that it is OK for a brand to have a website with inconsistent interfaces, I have come to the conclusion that it is more of a *mindset of tolerance* that has evolved over time. First, it is not uncommon to access a large organization's website that consists of a number of sub-sites created by different design groups and managed by different people in different departments. Second, during this research study, I found a number of websites with severe inconsistency problems

owned by well-known universities. Therefore, inconsistency might have become an expected or inevitable trait for all large-scale websites, which, therefore, does not have much to do with the quality of a specific brand.

5.5.2. Visual Guidance Affects Brand Image with Remarkable Gender Difference

Compared to consistency, visual guidance showed very strong influences on brand image, although this result was only true for females. Females who used the poor visual guidance websites gave significantly lower ratings to 13 out of the 22 brand identity attributes than those who used good visual guidance websites. To summarize, the female users thought that the university had a lower quality in regard to teaching and research, was less reputed, and provided less support to its students. Also, they thought that the students of this university had less favorable personalities and less developed social skills. On the one hand, this demonstrated how pervasive the effect of visual guidance was on females. Almost every core aspect of the university was affected. On the other hand, it is worth noting that, although the ratings of these attributes decreased significantly, most of them were still above the neutral point (4 of the 7-point scale). Only the ratings of B1 (High-quality teaching) and B6 (High ranking) dropped to the negative side (i.e., < 4). Therefore, most of the time, poor website visual guidance may greatly influence the image of a university, but may not be so devastating that could completely reverse people's opinions of it. In contrast, males only gave significantly lower ratings to one attribute – the university has a unique identity (B4) – in the poor visual guidance conditions. Compared to the original site (C+V+), the cluttered pages in the poor visual guidance versions made the university less distinguishable from its counterparts.

This is not the first time that significant discrepancies have been found between females and males in general decision-making (Gilligan, 1982; Meyers-Levy, 1988) or in assessing a design (Cyr & Bonanni, 2005; Flavián, Gurrea-Sarasa, & Orús Sanclemente, 2011; Holbrook, 1986; Moss, Gunn, & Heller, 2006; Moss, Gunn, & Kubacki, 2008; Simon & Peppas, 2005). Males are more prone to pursuing “agentic goals” than females (Meyers-Levy, 1988). They put more emphasis on “self-assertion” (Meyers-Levy, 1988) and behave more pragmatically and idiosyncratically (Flavián, Gurrea-Sarasa, & Orús Sanclemente, 2011). In contrast, females tend to be more “associative” (Flavián, Gurrea-Sarasa, & Orús Sanclemente, 2011), which encourages them to pursue harmony and good interpersonal relationships (Meyers-Levy, 1988). This may explain why the gender

differences occurred in this study. As shown in Table 20, despite the statistical significance, the absolute values of nearly all of the path coefficients were higher in the models that represented females' brand perceptions than in those models that represented the males' brand perceptions. In addition, the number of negative comments from females increased substantially when the poor visual guidance websites were used as stimuli in the follow-up interviews. These are all evidence that females are more likely than males to associate the visual guidance quality of a university website to the quality of the university itself. In simpler terms, a good university should not only provide good guidance in the real life, but also in the virtual world. To the contrary, for males, website visual guidance was only a peripheral factor. It was not related to how well a university could teach, research, or support students' careers. Therefore, as indicated in the follow-up interviews, males based their judgments primarily on the pragmatic factors described in the content of the website, such as the size and history of the university and the funding that it received. Since the results show that the content of the experiment website was well-composed (as indicated by the content-only group), males might have marginalized more of the impact of visual guidance (i.e., as long as the university could actually provide the services it promised, a poorly-design website did not matter much).

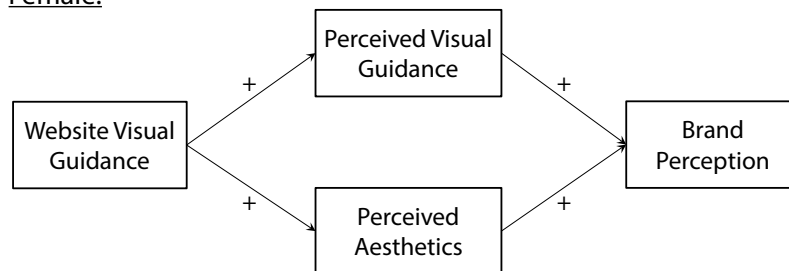
5.5.3. Toward a Gender-sensitive, Dual-mediation Model of Brand Perception

Given that a main cause of the gender differences is the additional effect of the experiment manipulation on visual guidance, it is important to understand the nature of this additional effect and why it did not occur to the males.

First, visual guidance has to do with the coloring, highlighting, and positioning of the webpage elements that could affect users' usability performances, rather than the attractiveness or fanciness of the design (i.e., aesthetics). However, when color contrasts were worsened or page elements were cluttered, it was inevitable that the pages would become less aesthetically pleasing. Second, as revealed by the follow-up interview, females emphasized the importance of aesthetical design for a university website much more than males. Third, according to the results presented in Chapter 4, visual guidance has a significant correlation with aesthetics. Its correlation was .62 and .50 with classical aesthetics and expressive aesthetics, respectively ($p < .01$). The values of these correlation coefficients are not too high ($< .08$), which guarantees that

they are different constructs. At the same time, the correlations again imply the possibility that any changes in visual guidance might affect perceived aesthetics. Finally, many prior studies have suggested that aesthetics could have a “halo effect” (De Angeli, Sutcliffe, & Hartmann, 2006; Dion, Berscheid, & Walster, 1972; Hartmann, Sutcliffe, & De Angeli, 2007; Meiners & Sheposh, 1977; Tractinsky, Katz, & Ikar, 2000): a beautiful person or artifact is considered good on every other aspect. For example, Tractinsky, Katz, and Ikar (2000) found strong evidence that the perceived usability of an ATM is determined more by aesthetical design than by its actual usability. De Angeli, Sutcliffe, and Hartmann (2006) provided a more elaborate picture that showed that the strength of the halo effect is dependent upon the seriousness of the context. The halo effect is strong in a casual or entertaining context. In a more serious scenario, such as when applying for a job, however, the aesthetics of a website become less important. Since the user tasks in this study are mainly targeted to occasional website visitors (instead of applicants of academic degrees or faculty positions), they fall more into the category of the casual context. Therefore, if the manipulations on visual guidance affects perceived aesthetics, perceived aesthetics may, in turn, have a halo effect on brand perceptions – a beautiful website represents a good brand.

Female:



Male (Might be spurious):

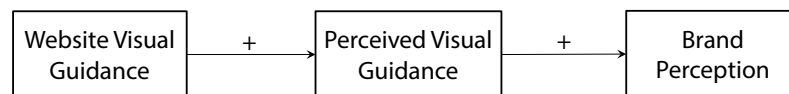


Figure 36. A gender-specific dual-mediation model for website visual guidance and brand perception.

Based on the foregoing analyses, I propose a gender-specific dual-mediation model (Figure 36). 1) For females, the effect of website visual guidance on brand perceptions is mediated by both perceived visual guidance and perceived aesthetics; and, 2) for males,

the effect of website visual guidance on brand perceptions is *probably* mediated only by perceived visual guidance. This mediation might be spurious because it cannot fully satisfy Baron and Kenny's criteria (section 5.4.5). The lack of the second mediation for males exists for two possible reasons. First, males did not perceive any change in aesthetics when visual guidance was worsened. Second, they perceived the aesthetics change, but did not take it into consideration when assessing the related brand. Based on the follow-up interviews, the second reason might be true because male users did complain about the aesthetics of the worsened experimental websites, but did not think that aesthetical design was critical for a university website. Of course, the model needs to be validated through further empirical studies.

5.5.4. Content Only Version Beat Full Website

The content only version was added to the experiment in order to gauge the quality of the website content. Also, I expected that the pure content document would yield equal branding effects to the original, high usability version (i.e., C+V+). Surprisingly, however, the former appeared to be better than the latter. This finding suggests that, as a brand unknown to the users, its website must be of very high quality to generate equal branding effects as a traditional non-interactive publication. It also implies that, although websites have already become a common tool for information consumption, a brochure or report could more transparently convey a rich set of articulated brand messages. This finding confirms the good practice of universities to add PDFs of their formal publications to complement their websites.

5.5.5. Understanding Interface Design for Branding: Synopsis of the Contribution

This study contributes to the advancement of the field of human-computer interaction in four aspects.

First, the limited effect of consistency on brand perceptions narrows down the research scope of online branding and relieves the tension of large organizations in regard to having inconsistent sub-sites. Slight interface inconsistencies will not cause trouble for the overall brand image. Instead of focusing on both inter-page navigation and intra-page design, future research may put more emphasis on the latter. In other words, researchers may concentrate on investigating what types of layout, highlighting, or element positioning could yield the best branding impact. As for large organizations, they

do not need to worry too much about hiring a different design team to create a website for a sub-institution. Even for organizations that now have a website with very different sub-sites, as long as each sub-site is well-designed, they do not need to allocate a large budget to standardize everything.

Table 24. A synopsis of the research findings (% Decrease: The percentage of the decrease in the brand perception rating when visual guidance was worsened).

Brand identity attributes sensitive to visual guidance for both genders	% Decrease
B4. The university has a <i>unique identity</i>	12.52
Brand identity attributes sensitive to visual guidance for females only	% Decrease
B1. The university provides <i>high-quality teaching</i>	38.26
B2. The university is committed to <i>academic excellence</i>	19.52
B3. The university has excellent teaching facilities	20.76
B6. The university has a <i>high ranking in US News</i>	23.15
B7. The university helps students to plan for their future	20.18
B9. The university offers a <i>personalizable path of study</i>	16.67
B10. They are self-disciplined	17.25
B11. They are <i>hard-working</i>	18.28
B12. They are <i>intelligent</i>	13.19
B13. They have excellent leadership	17.70
B14. They are highly involved in social activities	10.13
B15. They are the type of person whom I want to work with	19.96
Brand identity attributes tolerant to both visual guidance and consistency	
B5. The university has a <i>long history</i>	
B8. The campus of the university is a <i>safe place</i>	
B16. Deceitful – Sincere	
B17. Boring – Exciting	No
B18. Incompetent – Competent	Significant
B19. Low-class – Upper-class	Decrease
B20. Charmless – Charming	
B21. Passive – Active	
B22. Weak – Strong	

Second, when designing a new website, special attention needs to be paid to the brand identity attributes that females associated highly with website visual guidance. In Table 24, the “% Decrease” represents the sensitivity level of an attribute to visual guidance. The higher the sensitivity level, the more important it is to come up with a proper design. For example, with significantly worsened visual guidance, females’ perceptions of B1 (High-quality teaching) may decrease as much as 38%. If one or a set of webpages are going to be used to communicate this attribute to the users, then the designers should make sure that the visual guidance quality of these pages is satisfactory for females. In

particular, existing research suggests that females prefer websites that are designed by female designers than ones designed by male designers (Mirroring principle: Moss, Gunn, & Kubacki, 2008; Simon & Peppas, 2005). Therefore, it would be essential to involve more female designers, especially when creating brand-sensitive pages. Furthermore, in order to conduct a heuristic evaluation of an existing university website, in addition to regular inspection activities, evaluators may need to undertake an additional round of inspection to look specifically at the pages that convey the attributes sensitive to visual guidance. This practice may help evaluators spot problems overlooked in a regular usability evaluation.

Third, the dual-mediation model illustrates why worsened visual guidance has a different impact on the brand perception of females and males. This model spurs new research directions that can examine whether the second mediator for females' brand perceptions is really aesthetics as deduced from previous studies and as emerged from the follow-up interviews or some other attribute of which I am unaware. If it is indeed aesthetics, then a number of further questions could be raised, such as: Are males able to perceive the aesthetics change caused by visual guidance? If not, why not? In another situation, if males are able to perceive the aesthetics change, why do they not associate it with the image of the brand? In addition, the study results do not suggest significant differences in the effect of perceived visual guidance on brand perceptions between females and males; therefore, why is this effect not enough to significantly change males' brand perceptions? Is this effect spurious for males? Does high-quality content overshadow this effect? Since both consistency and visual guidance do not seem to affect males, is there any other website attribute besides content that males may associate with a university brand? If not, does that mean that males do not care about website design at all when assessing a brand? These questions have all been inspired by the dual-mediation model. By exploring these questions, website designers may get a better understanding of how to generate branding-effective designs for a specific gender group. Finally, this study offers a reusable set of methodologies and evaluation instruments to HCI researchers that can be used to replicate the same study, apply the study to a different website domain, or explore other important user experience dimensions that can affect brand perception.

5.5.6. Validity and Limits of the Study

Conducting user studies through Amazon Mechanical Turk has known advantages and limitations (Kittur, Chi, & Suh, 2008; Nelson & Stavrou, 2011). As for the advantages, it allowed me to recruit a large sample of participants who had diverse backgrounds at a relatively low cost. This ability not only increased the statistical power of the study, but also enabled me to go beyond the traditional recruiting of *local* university students (Nelson & Stavrou, 2011). Prospective students, faculty members, and occasional visitors were all part of the user profile. The main disadvantage of using Amazon Mechanical Turk was the cost of filtering out invalid responses and identifying participants who tried to game the system (e.g., getting the incentive without properly completing the tasks as instructed) (Nelson & Stavrou, 2011). For example, some participants gave up on all of the tasks or selected the same or random answers in the post-test questionnaire. In order to overcome this problem, I offered the participants a reasonable monetary incentive (\$2) for the 20- minute experiment. Then, I kept the number of instructions to the bare minimum and used simple and concise languages, which prevented the participants from intentionally skipping long and complicated instructions (Nelson & Stavrou, 2011). Next, I designed tasks that required a non-trivial effort to complete. In this way, it was easier for me to differentiate the time-on-task of an accurate session from a malicious respondent. Finally, multiple criteria were used to judge whether a response was valid (Kittur, Chi, & Suh, 2008), including time-on-task, task success rate, and statistical distributions.

The focus on university websites as one application domain manifests limits and advantages. On the one hand, the obvious limit is the external validity of the results, which may be restricted to websites and brands embodying the specific brand attributes that I investigated. On the other hand, the advantages are the strict control of potential confounding variables (e.g., using different domains, brand, and products), the reliance of a validated set of salient brand identity traits for the selected domain, and the strong reliability of the results.

Moreover, the use of a generic, fictitious brand, “A US State University” allowed me to explore the *pure* impact of website usability on the images of university brands in general, and avoided the possible influence of users’ prior knowledge about a brand. The picture could get much more complicated if well-established brands were used. In

particular, the users might be in favor of a university regardless of whether its website was good, which is in line with Jansen et al.'s research findings on users' biased opinions toward famous and non-famous search engines (Jansen, Zhang, & Schultz, 2009; Jansen, Zhang, & Zhang, 2007). This problem could be mitigated by involving a number of universities with different reputation levels. However, it may also introduce too many confounding variables, such as the locations of the universities (which has to do with the weather, economic conditions, and job opportunities), and make it too difficult to control for the variations in the two usability factors that I wanted to experiment on. Therefore, using a fictitious brand was considered a more appropriate approach.

Chapter 6. Summary of the Contributions

In summary, this dissertation contributes to the area of Human-Computer Interaction in the following ways.

First, the BREW evaluation approach enriches the existing HCI evaluation methods. The existing methods focus on gauging how usable an interactive product is and identifying the possible usability problems. These methods include lab-based and remote usability testing and heuristic evaluation. In contrast, the BREW approach supplements these methods with a method by which to assess how communicative a website is in transmitting key branding attributes to users and to identify the possible reasons why a certain attribute is not successfully communicated. The BREW approach provides rich guidelines by which to control for possible confounding variables, such as the users' prior knowledge about a brand and its website. In this way, researchers are allowed to see the *pure* effect of website visiting on brand attribute perception. At the same time, the approach offers enough flexibility for researchers to pick the evaluation elements that they need and customize their own user testing plan. In addition, with the support of the OFTEN field testing environment, researchers can either conduct the BREW evaluation in a lab or remotely online. Although no current counterpart exists for heuristic evaluation, the guidelines and criteria for creating branding-effective websites can be accumulated overtime by extensively applying the BREW approach to the evaluation of current company and institutional websites. By combining usability testing and branding effectiveness evaluation, HCI and eBranding professionals are empowered to create interactive products that are not only usable, but also communicative.

Second, the DEEP perceived usability questionnaire offers a brand new angle in regard to measuring and understanding perceived website usability. On the one hand, the questionnaire can directly reflect where the usability problem actually lies in a complex design. Capitalizing on the questionnaire, usability researchers may get quick, but very analytical feedback about the usability of a website by asking users to fill in the questionnaire without first conducting any formal usability tests. Typically, the website under evaluation may show a pop-up window on its interface to invite current users to fill in the DEEP questionnaire online after they have used the website for a while. In this way, the researchers may collect a large amount of responses in a short time at a low

cost. The existing questionnaires are less eligible for this role because they mostly capture holistic concepts, such as ease-of-use and learnability. They must be combined with a formal usability testing session to reveal why users consider a website difficult to learn and use.

On the other hand, the DEEP questionnaire shows that, among the five usability dimensions, the three most reliable and uncorrelated dimensions are: content, interface consistency, and visual guidance. The information architecture and navigation dimensions are highly correlated with interface consistency. This finding may imply that, in users' minds, website usability focuses on three things: whether the content is useful and easy to read, the cross-page navigation is consistent, and it is easy to find information on each page. This reveals the difference between usability experts and users in terms of how they understand and conceptualize website usability. It also shows that, in the future, if researchers are interested in investigating in-depth the impact of website usability on certain outcome measures, it might be more reasonable for them to center their experimental manipulations on those three aspects so that users can more accurately perceive the intended experimental effects.

Third, the controlled experiment and follow-up interview used to investigate the effect of interface consistency and visual guidance on brand perception offer an icebreaker in regard to demystifying the important issue of which design aspect of a website affects the communication of which specific brand attribute. First, with the methods and instruments offered by the controlled experiment, eBranding researchers can easily replicate or design similar research studies to further uncover the issue in different domains or research contexts (e.g., examining the issue in a more serious usage scenario). Second, the study results provide valuable insights in how the research scope of the issue could be greatly narrowed down in the future by excluding the non-significant impact of interface consistency. Third, the findings on the different effects of visual guidance on the communication of various brand attributes and the remarkable gender differences give web designers rich evidence on which problems to avoid in designing a university website and highlight the importance of designing for specific genders. The follow-up interview results not only confirm the findings from the controlled experiment, but also reveal why users are sensitive or insensitive to poor visual guidance when evaluating certain brand attributes and why males are not affected as

much by poor visual guidance. Finally, the hypothetical, gender-sensitive dual-mediation model opens up many new research directions, such as examining the possible mediator effect of aesthetics, testing the seemingly spurious effect of visual guidance on brand perception for males, and investigating males' tolerance levels to poor design qualities in judging the related brand image. Future research directions are discussed in more detail in the next chapter.

Chapter 7. Future Research Directions

The intersection of branding and interaction design is a broad, complex, and insufficiently explored area. It is difficult to separately manipulate the various components of an integral design artifact and even harder to ensure that users can accurately perceive experimental manipulations when allowed use the design artifact on their own. This dissertation took the initiative to decouple the various design dimensions of content-intensive websites through developing a design-oriented psychometric instrument (DEEP), devise a way to systematically capture users' brand perceptions based on existing branding models and HCI user testing techniques (BREW), and rigorously manipulate two of the most reliable website design dimensions to see their effects on users' opinions toward the typical university brand identity attributes. In spite of the fact that these research studies yielded rich outcomes in regard to revealing the proposed issue, they are just a snippet of the big picture when compared to the pervasive existence of branded interactive products. Therefore, this line of research can be further developed in, at least, the following directions.

Both the BREW evaluation approach and DEEP questionnaire can be further improved and validated. Both BREW and DEEP should be extensively applied to the evaluation of exiting content-intensive websites from various domains in order to test whether they are still reliable in gauging either the branding effectiveness or perceived usability of different web applications. In particular, through these further applications, eBranding researchers will be able to examine whether the suggested strategies in controlling prior brand knowledge are truly effective and whether the communication effectiveness of the tested websites can really be improved by taking recommendations from the evaluation findings. In addition, the extensive application of the BREW approach may yield rich knowledge on how to design branding-effective websites, which can be transformed into useful heuristics for web designers. The DEEP questionnaire, on the other hand, can be improved by either merging existing dimensions, which are highly correlated, or adding in additional dimensions, which are orthogonal to the existing dimensions.

The controlled experiment investigating the effects of interface consistency and visual guidance on brand perception can be extended in a number of ways. First, the experiment tested the effect based on fact-finding task scenarios. Whether the effects

would appear to be the same in a different type of task scenario is worth examining. For example, if users were asked to perform free-browsing tasks, the usability problems might become more noticeable because users do not need to focus their attention on finding the required information. Instead, they may simply lay back and take a look at every aspect of the website. In contrast, the effect can also be tested in more serious scenarios, such as making a payment, booking a flight, or applying for a job. Second, the effect can be investigated in a different domain area, such as commercial brands. As discussed in Chapter 5, the situation of a commercial brand can be much more complicated. It is a broad domain that covers a myriad amount of product and service categories (e.g., shoes, cars, housing). Researchers may pick a certain category to with which to begin. However, it might be difficult to anonymize an existing brand's website because the good websites might be easily recognized by users. Third, using fictitious brands may contaminate the ecological validity of the experiment findings. Therefore, another research direction might be to investigate real brands. Going this direction, researchers would need to find ways to strictly control for the effect of prior brand knowledge. In order to enhance the generalizability of the research findings, a sample of brands from a certain product category, instead of only one brand, should be investigated. Moreover, real company brands may cover a number of sub-brands. Researchers need to decide whether to experiment on the overall, company-level brand or a specific sub-brand. Finally, researchers may even jump out of the domain of websites to investigate the issue in another type of interactive application or product, such as a mobile device and application, video game, or driving technology.

Given that interface consistency does not have a significant effect on brand perception, researchers may focus on further investigating the effect of visual guidance. Moreover, we have already known that poor visual guidance may worsen brand perception. Future studies may explore how *good* design alternatives may generate different branding effects. For example, dropdown menus may yield different branding effects than ribbon menus. Dark color schemes might be more convincing than bright color schemes. Following this line, a variety of studies can be conducted in order to understand the relation between interface design and branding.

The gender-sensitive dual-mediation model may lead to three new research lines. First, it is certain that perceived visual guidance mediates the relation between website visual

guidance and female users' brand perceptions. The next step would be to investigate whether perceived aesthetics is the second mediator of the relation as I deduced from prior research findings. Second, the mediator effect of perceived visual guidance might be spurious for males because it did not lead to significant change in males' ratings on the brand identity attributes (although the effect appeared to be similar to the effect on the females). More studies need to be conducted in order to examine whether the effect was truly spurious or it was some other factors (perceived content might be a candidate) that stifled the effect. Third, males seemed to tolerate both types of poor design (inconsistency and poor visual guidance) in regard to judging a university's image. Future studies might investigate how tolerant males are to these factors by further worsening the problems and explore which usability factors can actually affect male users' brand perceptions.

Another perspective to interpret the gender difference is offered by Petty and Cacioppo (1986)'s Elaboration Likelihood Model. According to this theory, the amount of effort that users put in to process and evaluate information is affected by their "motivation" and "ability to think." When users are highly motivated and have strong ability to think, they tend to put more effort to reflect upon the information that they perceive and make more precise judgment. This process is defined as the "central route" to attitude change (Petty and Cacioppo, 1986). In contrast, when users are less motivated and have weaker ability to think, their attitude change tends to go through a "peripheral route," in which users put much less effort to process and evaluate the information that they perceive. According to the experiment and follow-up interview results, the usability problems in visual guidance seemed to affect females' opinions toward website usability through the central route, whereas the same problems seemed to affect males' opinions through the peripheral route. As suggested by Petty and Cacioppo (1986), both the central and peripheral route may have a strong impact on people's attitudes, which might explain why females and males perceived the visual guidance problems equally strong. However, within the central route, the elaborated thinking process might be the place where females built the relation between visual guidance design and brand image. In contrast, with the absence of the elaborated thinking process in the peripheral route, males did not put enough effort to establish the *design – brand image* relationship. Future research may come up with novel empirical experiments to investigate whether this hypothetical explanation is accurate.

In addition to interface consistency and visual guidance, researchers may also explore the influences of other website design factors on brand perception, such as content, information architecture, or navigation. Information architecture and navigation are highly correlated with interface consistency, so, presumably, they may not have a strong impact of brand perception. Still, empirical evidence is needed to justify this assumption. Finally, the content dimension is complex by itself. Narrowly speaking, it is about the way in which the text on a website is written. Broadly, it covers also the quality of the images, audios, and videos. Many research studies can be conduct only to investigate how these content-related attributes may affect a brand's image.

Appendices

Appendix A. Brand Perception Questionnaires Used in BREW Case Studies

Brand Perception Questionnaire for AT&T

Based on your experiences with the website, please fill out the following questionnaire regarding your opinions toward AT&T. It is very possible that you don't know every item listed below, but you can have a reasonable guess according to your impression of the website.

(Brand as Product)

What do you think of the products and services that AT&T provides?

1. The products that AT&T provides are innovative.

Strongly Disagree ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Strongly Agree

2. The products that AT&T provides are reliable.

Strongly Disagree ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Strongly Agree

3. The products that AT&T provides are high-quality.

Strongly Disagree ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Strongly Agree

4. The services that AT&T provides are innovative.

Strongly Disagree ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Strongly Agree

5. The services that AT&T provides are reliable.

Strongly Disagree ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Strongly Agree

6. The services that AT&T provides are high-quality.

Strongly Disagree ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Strongly Agree

7. The services that AT&T provides are highly international.

Strongly Disagree ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Strongly Agree

8. AT&T has the nation's largest 4G network.

Strongly Disagree ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Strongly Agree

9. AT&T operates the nation's largest Wi-Fi network.

Strongly Disagree ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Strongly Agree

10. AT&T U-verse provides a better TV experience.

Strongly Disagree ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Strongly Agree

11. AT&T provides excellent customer care.

Strongly Disagree ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Strongly Agree

(Brand as Organization)

What do you think of AT&T as an organization?

12. AT&T is committed to innovation.

Strongly Disagree ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Strongly Agree

13. AT&T understands what its customers want.

Strongly Disagree ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Strongly Agree

14. AT&T delivers what its customers want.

Strongly Disagree ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Strongly Agree

15. AT&T makes things easy for its customers.

Strongly Disagree ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Strongly Agree

16. AT&T empowers its customers to use technology safely.

Strongly Disagree ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Strongly Agree

17. AT&T empowers its customers to use technology securely.

Strongly Disagree ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Strongly Agree

18. AT&T is committed to do the right thing.

Strongly Disagree ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Strongly Agree

19. AT&T connects their users to the world in which they live and work.

Strongly Disagree ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Strongly Agree

(User Image)

Based on your experiences with the website, what do you think of the people who buy and use AT&T's products or services? Please give your rating to the following items.

The users of AT&T are:

Slow ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Fast

Ambitionless ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Ambitious

Self-doubting ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Confident

Wasting ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Saving

Unsocial ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Social

Old ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Young

(Brand as Person)

If you consider AT&T as a person, what is your opinion toward him/her? Please give your rating to the following items.

AT&T is:

Worthless ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Valuable

Unimaginative ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Creative

Neglecting ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Caring

Cautious ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Adventurous

Bad ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Righteous

Unreliable ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Reliable

Insecure ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Secure

Unsuccessful ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Successful

Follower ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Leader

Unhelpful ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Supportive

Inconsiderate ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Understanding

Unwise ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Intelligent

Brand Perception Questionnaire for Verizon

Based on your experiences with the website, please fill out the following questionnaire regarding your opinions toward Verizon. It is very possible that you don't know every item listed below, but you can have a reasonable guess according to your impression of the website.

(Brand as Product)

What do you think of the products and services that Verizon provides?

1. The products that Verizon provides are innovative.

Strongly Disagree ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Strongly Agree

2. The products that Verizon provides are reliable.

Strongly Disagree ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Strongly Agree

3. The products that Verizon provides are high-quality.

Strongly Disagree ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Strongly Agree

4. The services that Verizon provides are innovative.

Strongly Disagree ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Strongly Agree

5. The services that Verizon provides are reliable.

Strongly Disagree ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Strongly Agree

6. The services that Verizon provides are high-quality.

Strongly Disagree ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Strongly Agree

7. The services that Verizon provides are highly international.

Strongly Disagree ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Strongly Agree

8. Verizon has the nation's largest 4G network.

Strongly Disagree ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Strongly Agree

9. Verizon operates the nation's largest Wi-Fi network.

Strongly Disagree ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Strongly Agree

10. Verizon 's DIRECTV provides a better TV experience.

Strongly Disagree ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Strongly Agree

11. Verizon provides excellent customer care.

Strongly Disagree ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Strongly Agree

(Brand as Organization)

What do you think of Verizon as an organization?

12. Verizon is committed to innovation.

Strongly Disagree ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Strongly Agree

13. Verizon understands what its customers want.

Strongly Disagree ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Strongly Agree

14. Verizon delivers what its customers want.

Strongly Disagree ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Strongly Agree

15. Verizon makes things easy for its customers.

Strongly Disagree ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Strongly Agree

16. Verizon empowers its customers to use technology safely.

Strongly Disagree ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Strongly Agree

17. Verizon empowers its customers to use technology securely.

Strongly Disagree ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Strongly Agree

18. Verizon is committed to do the right thing.

Strongly Disagree ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Strongly Agree

19. Verizon connects their users to the world in which they live and work.

Strongly Disagree ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Strongly Agree

(User Image)

Based on your experiences with the website, what do you think of the people who buy and use Verizon's products or services? Please give your rating to the following items.

The users of Verizon are:

Slow ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Fast

Ambitionless ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Ambitious

Self-doubting ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Confident

Wasting ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Saving

Unsocial ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Social

Old ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Young

(Brand as Person)

If you consider Verizon as a person, what is your opinion toward him/her? Please give your rating to the following items.

Verizon is:

Worthless ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Valuable

Unimaginative ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Creative

Neglecting ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Caring

Cautious ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Adventurous

Bad ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Righteous

Unreliable ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Reliable

Insecure ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Secure

Unsuccessful ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Successful

Follower ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Leader

Unhelpful ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Supportive

Inconsiderate ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Understanding

Unwise ----- 1 ----- 2 ----- 3 ----- 4 (I don't know/neutral) ----- 5 ----- 6 ----- 7 ----- Intelligent

Appendix B. User Tasks for the BREW Case Studies

User Tasks – AT&T

Task#1. Suppose you are interested in receiving digital TV in your home and you want to record some of your favorite TV shows. Please check out up to how many hours of HD video content you can record with the U-verse DVR.

Task#2. You want to get a cheap, basic internet service for occasional use in your apartment. Please check out what the downstream speed is for a basic DSL service.

Task#3. Your best friend went to study in another state. You want to add an unlimited long distance plan to your home phone so that you can call him/her regularly. How much is the plan?

Task#4. You are planning to buy a new Samsung Galaxy S 4 (16GB). Please check what special offer you can get.

User Tasks – Verizon

Task#1. Suppose you are interested in receiving digital TV in your home and you want to record some of your favorite TV shows. Please check out up to how many hours of SD video content you can record with the DIRECTV DVR receiver.

Task#2. You want to get a cheap, basic internet service for occasional use in your apartment. Please check out what the download speed is for the basic high speed internet service.

Task#3. Your best friend went to study in another state. You want to add an unlimited long distance plan to your home phone so that you can call him/her regularly. Please check what special offer you can get.

Task#4. You are planning to buy a new Windows Phone 8X by HTC. Please check the price.

Appendix C. Demographics Questions

Thanks again! You are almost done.

Please fill in the following questions about yourself and click Submit to get your reward code.

1. Sex
 - ☐ Male
 - ☐ Female
2. Age
 - ☐ 17 or below
 - ☐ 18 to 20
 - ☐ 21 to 23
 - ☐ 24 to 26
 - ☐ 27 to 30
 - ☐ 31 to 40
 - ☐ 41 or older
3. The highest level of education that you have completed
 - ☐ Lower than high school
 - ☐ High school diploma
 - ☐ First-year college
 - ☐ Second-year college
 - ☐ Third-year college
 - ☐ Bachelor degree
 - ☐ Master's degree
 - ☐ Doctoral degree

Appendix D. No Significant Difference Was Found Between AT&T and Verizon in the Perceptions of All Brand Identity Attributes after In-depth Website Exposure

	AT&T	Verizon	Difference	Significance
Innovative Products	5.06	5.12	0.06	Non
Reliable Products	5.12	5.38	0.26	Non
High-Quality Products	5.52	5.76	0.24	Non
Innovative Services	4.66	4.98	0.32	Non
Reliable Services	4.96	5.05	0.09	Non
High-Quality Services	5.08	5.40	0.32	Non
Highly International Services	4.76	4.76	0.00	Non
Has the nation's largest 4G network	4.44	5.05	0.61	Non
Operates the nation's largest Wi-Fi network	4.32	4.67	0.35	Non
Provides a better TV experience	4.62	4.67	0.05	Non
Provides excellent customer care	4.64	4.69	0.05	Non
Committed to innovation	4.90	5.02	0.12	Non
Understands what its customers want	4.90	4.83	-0.07	Non
Delivers what its customers want	4.88	4.95	0.07	Non
Makes things easy for its customers	4.34	4.76	0.42	Non
Empowers its customers to use technology safely	4.84	4.93	0.09	Non
Empowers its customers to use technology securely	4.88	5.05	0.17	Non
Committed to do the right thing	4.74	4.83	0.09	Non
Connects their users to the world in which they live and work	5.34	5.69	0.35	Non
Slow-Fast	4.72	5.10	0.38	Non
Ambitionless-Ambitious	4.48	5.05	0.57	Non
Self-doubting-Confident	4.58	5.14	0.56	Non
Wasting-Saving	4.24	4.50	0.26	Non
Old-Young	4.24	4.81	0.57	Non
Worthless-Valuable	4.54	4.88	0.34	Non
Unimaginative-Creative	4.46	4.83	0.37	Non
Neglecting-Caring	4.06	4.43	0.37	Non
Cautious-Adventurous	4.64	4.74	0.10	Non
Bad-Righteous	4.22	4.33	0.11	Non
Unreliable-Reliable	4.42	4.95	0.53	Non
Insecure-Secure	4.68	4.93	0.25	Non
Unsuccessful-Successful	5.10	5.45	0.35	Non
Follower-Leader	4.62	4.88	0.26	Non
Unhelpful-Supportive	4.26	4.81	0.55	Non
Inconsiderate-Understanding	4.46	4.60	0.14	Non
Unwise-Intelligent	4.78	5.14	0.36	Non

*Data from high brand familiarity group

Appendix E. Candidate Measurement Items Selected from Existing Questionnaires (49 Items)

*The source where a question is selected from is annotated in the bracket.

Perceived Content

Does it provide clarity of wording? (PUTQ)

Are long data items partitioned? (PUTQ)

The information provided with this system was clear. (IBM)

The information was effective in helping me complete the tasks and scenarios. (IBM)

The organization of information on the system screen was clear. (IBM)

This system has all the functions and capabilities I expect it to have. (IBM)

Everything on this web site is easy to understand. (WAMMI)

Perceived Navigation

It requires the fewest steps possible to accomplish what I want to do with it. (USE)

Is the coding consistent across displays, menu options? (PUTQ)

Is the label location consistent? (PUTQ)

Is the guidance information always available? (PUTQ)

Does it indicate current position in menu structure? (PUTQ)

Provide shortcuts. (PHUE)

Give the user a way to review / return-to previous contexts. (PHUE)

It was easy to find the information I needed. (IBM)

I can quickly find what I want on this web site. (WAMMI)

Going from one part to another is easy on this web site. (WAMMI)

Remembering where I am on this web site is easy/difficult. (WAMMI)

The return to higher-level pages requires only one simple key action. (PUTQ)

It is difficult to move around this web site. (WAMMI)

This web site helps me find what I am looking for. (WAMMI)

Learning to find my way around this web site is a problem. (WAMMI)

Perceived Labeling

I can use it without written instructions. (USE)

I don't notice any inconsistencies as I use it. (USE)

Is the wording familiar? (PUTQ)

Is the label format consistent? (PUTQ)

Is the wording consistent across displays? (PUTQ)

Use of terms throughout system (inconsistent/consistent) (QUIS)

I get what I expect when I click on things on this web site. (WAMMI)

Perceived Cognitive Effort

It saves me time when I use it. (USE)

Using it is effortless. (USE)

I can recover from mistakes quickly and easily. (USE)

I learned to use it quickly. (USE)

I easily remember how to use it. (USE)

It is easy to learn to use it. (USE)

Is the screen density reasonable? (PUTQ)

Are groups of information demarcated? (PUTQ)

Information should be in a logical, natural order. (Organization of information) (PHUE)

Give the user a way to preview where to go, what will happen. (PHUE)

This web site seems logical to me. (WAMMI)

Perceived Graphics

Is the assignment of colour codes conventional? (PUTQ)

Does it provide easily distinguished colours? (PUTQ)

Reading characters on the screen (hard/easy). (QUIS)

Highlighting simplifies task (not at all/very much). (QUIS)

The pages on this web site are very attractive. (WAMMI)

This web site has some annoying features. (WAMMI)

Perceived Technological Performance

This web site is too slow. (WAMMI)

I feel efficient when I'm using this web site. (WAMMI)

Make it difficult to make errors. (PHUE)

Candidate Measurement Items Selected from Existing Questionnaires (49 Items)

Questionnaire	Total	Content	Navigation	Labeling	Cognitive	Graphics	Performance
PUTQ	14	2	5	3	2	2	0
IBM	5	4	1	0	0	0	0
WAMMI	13	1	6	1	1	2	2
USE	9	0	1	2	6	0	0
PHUE	5	0	2	0	2	0	1
QUIS	3	0	0	1	0	2	0
Total	49	7	15	7	11	6	3

Appendix F. Initial Questionnaire Created for Internal Rating

*** Items with minor or no change from the original version are in bold.**

**** The actual questionnaire used in internal rating was not numbered.**

The following statements are used for capturing users' perceived usability of websites. Please ignore whether a statement is positive or negative. Just give a ranking of the statements under each dimension, based on the importance of the issue that the statement touches.

Perceived Content

___ (1) The wording of the text was clear. (PUTQ)

___ (2) The content (including text, pictures, audios, and videos etc.) was easy to understand. (WAMMI)

___ (3) The content (including text, pictures, audios, and videos etc.) was what I expected. (IBM)

___ (4) The content (including text, pictures, audios, and videos etc.) did not contain errors or misspellings. (MiLE+)

___ (5) There were many grammatical errors and typos on the website. (MiLE+)

___ (6) The content was up to date. (MiLE+)

___ (7) The content looked trustworthy. (MiLE+)

___ (8) The text was concise and easy to read. (MiLE+)

___ (9) There were many irrelevant pictures on the website. (MiLE+)

___ (10) There were many irrelevant videos on the website. (MiLE+)

Perceived Structure and Information Architecture

___ (1) The long text on the website was well partitioned into sections. (PUTQ)

___ (2) Related web pages were scattered throughout the website. (MiLE+)

___ (3) I could easily locate all related information in one section. (MiLE+)

- ___ (4) I could quickly get to know the structure of the website by skimming its homepage. (MiLE+)
- ___ (5) The organization of the website was clear. (PHUE)
- ___ (6) Under each section of the website, the web pages were well organized. (MiLE+)
- ___ (7) The links on each web page were well organized. (MiLE+)
- ___ (8) Some web pages were too long. (MiLE+)

Perceived Navigation

- ___ (1) It required only a few steps to accomplish the tasks. (USE)
- ___ (2) The website provided enough guidance for me to navigate through the content. (PUTQ)
- ___ (3) It was easy for me to return to previous pages. (PUTQ)
- ___ (4) **It was easy to find the information I needed on the website. (IBM)**
- ___ (5) **Going from one section to another was easy on this website. (WAMMI)**
- ___ (6) **Remembering where I was on this website was difficult. (WAMMI)**
- ___ (7) Returning to higher-level pages was immediate. (PUTQ)
- ___ (8) **It was difficult for me to move around this website. (WAMMI)**
- ___ (9) **This website helped me find what I was looking for. (WAMMI)**
- ___ (10) **Learning to find my way around this website was a problem. (WAMMI)**

Perceived Labeling

- ___ (1) The meanings of the main navigation labels were straightforward. (MiLE+)
- ___ (2) When I skimmed through the labels I confidently knew where to click. (MiLE+)
- ___ (3) The links on the website were clear. (MiLE+)
- ___ (4) The headings of the web pages were easy to understand. (MiLE+)
- ___ (5) The captions of the pictures were clear. (MiLE+)
- ___ (6) The captions of the videos were clear. (MiLE+)
- ___ (7) **I got what I expected when I clicked on things on this website. (WAMMI)**

Perceived Cognitive Effort

- ___ (1) **Using this website was effortless. (USE)**
- ___ (2) Using this website made me feel tired. (MiLE+)
- ___ (3) **I learned to use this website quickly. (USE)**
- ___ (4) The web pages were dense. (PUTQ)

Perceived Graphics

- ___ (1) The colors used on the website made me feel uncomfortable. (MiLE+)
- ___ (2) The colors helped me to distinguish different sections of the website. (MiLE+)
- ___ (3) The highlighted areas of a page helped me locate the information I needed. (QUIS)
- ___ (4) **The pages on this website were very attractive. (WAMMI)**
- ___ (5) This website had some annoying colors. (MiLE+)
- ___ (6) The text on the website was big enough to read. (MiLE+)
- ___ (7) The colors of the text made it easy to read. (MiLE+)
- ___ (8) The fonts used on this website were strange. (MiLE+)
- ___ (9) The home page of this website was pleasant. (MiLE+)
- ___ (10) The layout of the pages throughout the website was consistent. (MiLE+)
- ___ (11) I noticed abrupt changes in the layout of the pages. (MiLE+)

Perceived Technological Performance

- ___ (1) The pages took too long to load. (WAMMI)
- ___ (2) I needed to install plug-ins to view some of the content. (MiLE+)
- ___ (3) When I made a mistake, the website promptly helped me to recover. (PHUE)
- ___ (4) Multimedia (audio and video) functioned well on this website. (MiLE+)

Details about the Initial Usability Questionnaire Developed for Internal Rating

Questionnaire	Total	Cont	Strut	Nav	Lab	Cog	Grap	Tech
PUTQ	6	1	1	3	0	1	0	0
IBM	2	1	0	1	0	0	0	0
WAMMI	9	1	0	5	1	0	1	1
USE	3	0	0	1	0	2	0	0
PHUE	2	0	1	0	0	0	0	1
QUIS	1	0	0	0	0	0	1	0
MiLE+	31	7	6	0	6	1	9	2
Total	54	10	8	10	7	4	11	4

*Cont = Content; Strut = Structure and Information Architecture; Nav = Navigation; Lab = Labeling; Cog = Cognitive Effort; Grap = Graphics; Tech = Technological Performance.

Appendix G. Perceived Usability Questionnaire Used in Online Field Test 1

* Items with minor or no change from the original version are in bold.

** SD = Strongly Disagree; D = Disagree; N = Neutral; A = Agree; SA = Strongly Agree; N/A = Not Applicable

Perceived Content

1. The wording of the text was clear. (PUTQ)	SD D N A SA N/A
2. The content (including text, pictures, audios, and videos etc.) was easy to understand. (WAMMI)	SD D N A SA N/A
3. The content (including text, pictures, audios, and videos etc.) was what I expected. (IBM)	SD D N A SA N/A
4. The content (including text, pictures, audios, and videos etc.) did not contain errors or misspellings. (MiLE+)	SD D N A SA N/A
5. There were many grammatical errors and typos on the website. (MiLE+)	SD D N A SA N/A
6. The content was up to date. (MiLE+)	SD D N A SA N/A
7. The content looked trustworthy. (MiLE+)	SD D N A SA N/A
8. The text was concise. (MiLE+)	SD D N A SA N/A
9. The text was easy to read. (MiLE+)	SD D N A SA N/A

Perceived Structure and Information Architecture

10. The long text on the website was well partitioned into sections. (PUTQ)	SD D N A SA N/A
11. Related web pages were scattered throughout the website. (MiLE+)	SD D N A SA N/A
12. I could easily locate all related information in one section. (MiLE+)	SD D N A SA N/A
13. I could quickly get to know the structure of the website by skimming its homepage. (MiLE+)	SD D N A SA N/A
14. The organization of the website was clear. (PHUE)	SD D N A SA N/A
15. Under each section of the website, the web pages were well organized. (MiLE+)	SD D N A SA N/A

16. The links on each web page were well organized. (MiLE+)	SD D N A SA N/A
17. Some web pages were too long. (MiLE+)	SD D N A SA N/A

Perceived Navigation

18. It required only a few steps to accomplish the tasks. (USE)	SD D N A SA N/A
19. The website provided enough guidance for me to navigate through the content. (PUTQ)	SD D N A SA N/A
20. It was easy for me to return to previous pages. (PUTQ)	SD D N A SA N/A
21. It was easy to find the information I needed on the website. (IBM)	SD D N A SA N/A
22. Going from one section to another was easy on this website. (WAMMI)	SD D N A SA N/A
23. Remembering where I was on this website was difficult. (WAMMI)	SD D N A SA N/A
24. Returning to higher-level pages was immediate. (PUTQ)	SD D N A SA N/A
25. It was difficult for me to move around this website. (WAMMI)	SD D N A SA N/A
26. This website helped me find what I was looking for. (WAMMI)	SD D N A SA N/A
27. Learning to find my way around this website was a problem. (WAMMI)	SD D N A SA N/A

Perceived Labeling

28. The meanings of the main navigation labels were straightforward. (MiLE+)	SD D N A SA N/A
29. When I skimmed through the labels I confidently knew where to click. (MiLE+)	SD D N A SA N/A
30. The links on the website were clear. (MiLE+)	SD D N A SA N/A
31. The headings of the web pages were easy to understand. (MiLE+)	SD D N A SA N/A
32. The captions of the pictures were clear. (MiLE+)	SD D N A SA N/A
33. The captions of the videos were clear. (MiLE+)	SD D N A SA N/A

34. I got what I expected when I clicked on things on this website. (WAMMI)	SD D N A SA N/A
--	-----------------

Perceived Cognitive Effort

35. Using this website was effortless. (USE)	SD D N A SA N/A
36. Using this website made me feel tired. (MiLE+)	SD D N A SA N/A
37. I learned to use this website quickly. (USE)	SD D N A SA N/A
38. The web pages were dense. (PUTQ)	SD D N A SA N/A

Perceived Graphics

39. The colors used on the website made me feel uncomfortable. (MiLE+)	SD D N A SA N/A
40. The colors helped me to distinguish different sections of the website. (MiLE+)	SD D N A SA N/A
41. The highlighted areas of a page helped me locate the information I needed. (QUIS)	SD D N A SA N/A
42. The pages on this website were pleasant to look at. (WAMMI)	SD D N A SA N/A
43. This website had some annoying colors. (MiLE+)	SD D N A SA N/A
44. The text on the website was big enough to read. (MiLE+)	SD D N A SA N/A
45. The text on the website was readable. (MiLE+)	SD D N A SA N/A
46. The colors of the text made it easy to read. (MiLE+)	SD D N A SA N/A
47. The home page of this website was pleasant. (MiLE+)	SD D N A SA N/A
48. The layout of the pages throughout the website was consistent. (MiLE+)	SD D N A SA N/A
49. I noticed abrupt changes in the layout of the pages. (MiLE+)	SD D N A SA N/A

Perceived Technological Performance

50. The pages took too long to load. (WAMMI)	SD D N A SA N/A
51. I needed to install plug-ins to view some of the content. (MiLE+)	SD D N A SA N/A

52. When I made a mistake, the website promptly helped me to recover. (PHUE)	SD	D	N	A	SA	N/A
53. Multimedia (audio and video) functioned well on this website. (MiLE+)	SD	D	N	A	SA	N/A

Details about Perceived Usability Questionnaire Used in the Online Field Test 1

Questionnaire	Total	Cont	Strut	Nav	Lab	Cog	Grap	Tech
PUTQ	6	1	1	3	0	1	0	0
IBM	2	1	0	1	0	0	0	0
WAMMI	9	1	0	5	1	0	1	1
USE	3	0	0	1	0	2	0	0
PHUE	2	0	1	0	0	0	0	1
QUIS	1	0	0	0	0	0	1	0
MiLE+	30	6	6	0	6	1	9	2
Total	53	9	8	10	7	4	11	4

*Cont = Content; Strut = Structure and Information Architecture; Nav = Navigation; Lab = Labeling; Cog = Cognitive Effort; Grap = Graphics; Tech = Technological Performance.

Appendix H. Correlations of the Nine Usability Dimensions Extracted from EFA

		Perceived Content	Perceived Structure and Information Architecture	Perceived Navigation	Perceived Labeling	Perceived Cognitive Effort	Perceived Graphics_1	Perceived Graphics_2	Perceived Graphics_3	Perceived Graphics_4
Perceived Content	Pearson Correlation Sig. (2-tailed) N	1 0 196	.661** 0 196	.564** 0 196	.614** 0 196	.548** 0 196	.536** 0 196	.447** 0 196	.403** 0 196	.326** 0 196
Perceived Structure and Information Architecture	Pearson Correlation Sig. (2-tailed) N	.661** 0 196	1 0 196	.842** 0 196	.815** 0 196	.795** 0 196	.385** 0 196	.531** 0 196	.518** 0 196	.304** 0 196
Perceived Navigation	Pearson Correlation Sig. (2-tailed) N	.564** 0 196	.842** 0 196	1 0 196	.820** 0 196	.875** 0 196	.316** 0 196	.411** 0 196	.433** 0 196	.224** 0.002 196
Perceived Labeling	Pearson Correlation Sig. (2-tailed) N	.614** 0 196	.815** 0 196	.820** 0 196	1 0 196	.829** 0 196	.396** 0 196	.468** 0 196	.525** 0 196	.235** 0.001 196
Perceived Cognitive Effort	Pearson Correlation Sig. (2-tailed) N	.548** 0 196	.795** 0 196	.875** 0 196	.829** 0 196	1 0 196	.332** 0 196	.401** 0 196	.446** 0 196	.291** 0 196
Perceived Graphics_1	Pearson Correlation Sig. (2-tailed) N	.536** 0 196	.385** 0 196	.316** 0 196	.396** 0 196	.332** 0 196	1 0 196	.268** 0 196	.317** 0 196	.192** 0.007 196
Perceived Graphics_2	Pearson Correlation Sig. (2-tailed) N	.447** 0 196	.531** 0 196	.411** 0 196	.468** 0 196	.401** 0 196	.268** 0 196	1 0 196	.301** 0 196	.297** 0 196
Perceived Graphics_3	Pearson Correlation Sig. (2-tailed) N	.403** 0 196	.518** 0 196	.433** 0 196	.525** 0 196	.446** 0 196	.317** 0 196	.301** 0 196	1 0 196	.210** 0.003 196
Perceived Graphics_4	Pearson Correlation Sig. (2-tailed) N	.326** 0 196	.304** 0 196	.224** 0.002 196	.235** 0.001 196	.291** 0 196	.192** 0.007 196	.297** 0 196	.210** 0.003 196	1 196

** Correlation is significant at the 0.01 level (2-tailed).

Appendix I. Correlations among the Usability Dimensions and Success Rate after CFA

		Perceived Content	Perceived Structure and Information Architecture	Perceived Navigation	Perceived Cognitive Effort	Perceived Layout Consistency	Perceived Visual Guidance	Perceived Usability	Task Success Rate (Performance Based Usability)
Perceived Content	Pearson Correlation	1	.611**	.571**	.540**	.423**	.394**	.768**	.243**
	Sig. (2-tailed)		0	0	0	0	0	0	0.001
	N	196	196	196	196	196	196	196	196
Perceived Structure and Information Architecture	Pearson Correlation	.611**	1	.807**	.768**	.505**	.481**	.899**	.299**
	Sig. (2-tailed)	0		0	0	0	0	0	0
	N	196	196	196	196	196	196	196	196
Perceived Navigation	Pearson Correlation	.571**	.807**	1	.869**	.396**	.429**	.888**	.465**
	Sig. (2-tailed)	0	0		0	0	0	0	0
	N	196	196	196	196	196	196	196	196
Perceived Cognitive Effort	Pearson Correlation	.540**	.768**	.869**	1	.401**	.446**	.868**	.338**
	Sig. (2-tailed)	0	0	0		0	0	0	0
	N	196	196	196	196	196	196	196	196
Perceived Layout Consistency	Pearson Correlation	.423**	.505**	.396**	.401**	1	.301**	.608**	0.121
	Sig. (2-tailed)	0	0	0	0		0	0	0.091
	N	196	196	196	196	196	196	196	196
Perceived Visual Guidance	Pearson Correlation	.394**	.481**	.429**	.446**	.301**	1	.617**	.153*
	Sig. (2-tailed)	0	0	0	0	0		0	0.032
	N	196	196	196	196	196	196	196	196
Perceived Usability	Pearson Correlation	.768**	.899**	.888**	.868**	.608**	.617**	1	.368**
	Sig. (2-tailed)	0	0	0	0	0	0		0
	N	196	196	196	196	196	196	196	196
Task Success Rate (Performance Based Usability)	Pearson Correlation	.243**	.299**	.465**	.338**	0.121	.153*	.368**	1
	Sig. (2-tailed)	0.001	0	0	0	0.091	0.032	0	
	N	196	196	196	196	196	196	196	196

** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).

Appendix J. Perceived Usability Questionnaire Used in Online Field Test 2

* Items with minor or no change from the original version are in bold.

** Items kept from Test 1 are labeled by their item numbers in the former version.

*** Newly added items are labeled as (New).

**** SD = Strongly Disagree; D = Disagree; N = Neutral; A = Agree; SA = Strongly Agree; N/A = Not Applicable

Perceived Content

The wording of the text was clear (1) (PUTQ).	SD D N A SA N/A
The content (including text, pictures, audios, and videos etc.) was easy to understand (2) (WAMMI).	SD D N A SA N/A
The text was concise (8) (MiLE+).	SD D N A SA N/A
The text was easy to read (9) (MiLE+).	SD D N A SA N/A
The content was what I wanted (New) (MiLE+).	SD D N A SA N/A
The text was useful (New) (MiLE+).	SD D N A SA N/A
The text was relevant (New) (MiLE+).	SD D N A SA N/A

Perceived Structure and Information Architecture

I could quickly get to know the structure of the website by skimming its homepage (13) (MiLE+).	SD D N A SA N/A
The organization of the website was clear (14) (PHUE).	SD D N A SA N/A
Under each section of the website, the web pages were well organized (15) (MiLE+).	SD D N A SA N/A

Perceived Navigation

It was easy to find the information I needed on the website (21) (IBM).	SD D N A SA N/A
This website helped me find what I was looking for (26) (WAMMI).	SD D N A SA N/A
Learning to find my way around this website was a problem (27) (WAMMI).	SD D N A SA N/A

I got what I expected when I clicked on things on this website (34) (WAMMI).	SD D N A SA N/A
---	-----------------

Perceived Cognitive Effort

Using this website was effortless (35) (USE).	SD D N A SA N/A
Using this website made me feel tired (36) (MiLE+).	SD D N A SA N/A
I learned to use this website quickly (37) (USE).	SD D N A SA N/A

Perceived Layout Consistency

The layout of the pages throughout the website was consistent (48) (MiLE+).	SD D N A SA N/A
I noticed abrupt changes in the layout of the pages (49) (MiLE+).	SD D N A SA N/A
The layout of the pages was pleasant (New) (MiLE+).	SD D N A SA N/A
The abrupt layout changes made me lose my way (New) (MiLE+).	SD D N A SA N/A
The layout under each section of the website was consistent (New) (MiLE+).	SD D N A SA N/A

Perceived Visual Guidance

The colors helped me to distinguish different sections of the website (40) (MiLE+).	SD D N A SA N/A
The highlighted areas of a page helped me locate the information I needed (41) (QUIS).	SD D N A SA N/A
I got to know the content of a page by skimming the highlighted areas (New) (MiLE+).	SD D N A SA N/A
Some of the highlighted areas distracted me from locating the information I needed (New) (MiLE+).	SD D N A SA N/A

Details about Perceived Usability Questionnaire Used in Online Field Test 2

Questionnaire	Total	Cont	Strut	Nav	Cog	Lay	Vis
PUTQ	1	1	0	0	0	0	0
IBM	1	0	0	1	0	0	0
WAMMI	4	1	0	3	0	0	0
USE	2	0	0	0	2	0	0
PHUE	1	0	1	0	0	0	0
QUIS	1	0	0	0	0	0	1
MILE+	16	5	2	0	1	5	3
Total	26	7	3	4	3	5	4

*Cont = Content; Strut = Structure and Information Architecture; Nav = Navigation; Cog = Cognitive Effort; Lay = Layout Consistency; Vis = Visual Guidance.

Appendix K. DEEP: Design-oriented Evaluation of Perceived Usability

* Items with minor or no change from the original version are in bold.

** SD = Strongly Disagree; D = Disagree; N = Neutral; A = Agree; SA = Strongly Agree; N/A = Not Applicable

Perceived Content

The wording of the text was clear (PUTQ).	SD D N A SA N/A
The content (including text, pictures, audios, and videos etc.) was easy to understand (WAMMI).	SD D N A SA N/A
The text was useful (MiLE+).	SD D N A SA N/A
The text was relevant (MiLE+).	SD D N A SA N/A

Perceived Structure and Information Architecture

I could quickly get to know the structure of the website by skimming its homepage (MiLE+).	SD D N A SA N/A
The organization of the website was clear (PHUE).	SD D N A SA N/A
Under each section of the website, the web pages were well organized (MiLE+).	SD D N A SA N/A

Perceived Navigation

It was easy to find the information I needed on the website (IBM).	SD D N A SA N/A
This website helped me find what I was looking for (WAMMI).	SD D N A SA N/A
I got what I expected when I clicked on things on this website (WAMMI).	SD D N A SA N/A

Perceived Cognitive Effort

Using this website was effortless (USE).	SD D N A SA N/A
Using this website made me feel tired (MiLE+).	SD D N A SA N/A
I learned to use this website quickly (USE).	SD D N A SA N/A

Perceived Layout Consistency

The layout of the pages throughout the website was consistent (MiLE+).	SD D N A SA N/A
I noticed abrupt changes in the layout of the pages (MiLE+).	SD D N A SA N/A
The layout under each section of the website was consistent (MiLE+).	SD D N A SA N/A

Perceived Visual Guidance

The colors helped me to distinguish different sections of the website (MiLE+).	SD D N A SA N/A
The highlighted areas of a page helped me locate the information I needed (QUIS).	SD D N A SA N/A
I got to know the content of a page by skimming the highlighted areas (MiLE+).	SD D N A SA N/A

Details about the DEEP Questionnaire

Questionnaire	Total	Cont	Strut	Nav	Cog	Lay	Vis
PUTQ	1	1	0	0	0	0	0
IBM	1	0	0	1	0	0	0
WAMMI	3	1	0	2	0	0	0
USE	2	0	0	0	2	0	0
PHUE	1	0	1	0	0	0	0
QUIS	1	0	0	0	0	0	1
MiLE+	10	2	2	0	1	3	2
Total	19	4	3	3	3	3	3

*Cont = Content; Strut = Structure and Information Architecture; Nav = Navigation; Cog = Cognitive Effort; Lay = Layout Consistency; Vis = Visual Guidance.

The final version of DEEP is also available online as web form at:

<http://discern.uits.iu.edu:8670/DEEP/deep.html>

Appendix L. Correlations among Perceived Usability Dimensions, Overall Perceived Usability and Task Success Rate

		Perceived Content	Perceived Structure and Information Architecture	Perceived Navigation	Perceived Cognitive Effort	Perceived Layout Consistency	Perceived Visual Guidance	Task Success Rate (Performance Based Usability)	Perceived Usability
Perceived Content	Pearson Correlation	1	.715**	.663**	.614**	.475**	.536**	.424**	.793**
	Sig. (2-tailed)	0	0	0	0	0	0	0	0
	N	362	362	362	362	362	362	362	362
Perceived Structure and Information Architecture	Pearson Correlation	.715**	1	.841**	.806**	.582**	.673**	.494**	.919**
	Sig. (2-tailed)	0	0	0	0	0	0	0	0
	N	362	362	362	362	362	362	362	362
Perceived Navigation	Pearson Correlation	.663**	.841**	1	.846**	.601**	.668**	.570**	.918**
	Sig. (2-tailed)	0	0	0	0	0	0	0	0
	N	362	362	362	362	362	362	362	362
Perceived Cognitive Effort	Pearson Correlation	.614**	.806**	.846**	1	.531**	.692**	.554**	.891**
	Sig. (2-tailed)	0	0	0	0	0	0	0	0
	N	362	362	362	362	362	362	362	362
Perceived Layout Consistency	Pearson Correlation	.475**	.582**	.601**	.531**	1	.486**	.286**	.721**
	Sig. (2-tailed)	0	0	0	0	0	0	0	0
	N	362	362	362	362	362	362	362	362
Perceived Visual Guidance	Pearson Correlation	.536**	.673**	.668**	.692**	.486**	1	.368**	.803**
	Sig. (2-tailed)	0	0	0	0	0	0	0	0
	N	362	362	362	362	362	362	362	362
Task Success Rate (Performance Based Usability)	Pearson Correlation	.424**	.494**	.570**	.554**	.286**	.368**	1	.537**
	Sig. (2-tailed)	0	0	0	0	0	0	0	0
	N	362	362	362	362	362	362	362	362
Perceived Usability	Pearson Correlation	.793**	.919**	.918**	.891**	.721**	.803**	.537**	1
	Sig. (2-tailed)	0	0	0	0	0	0	0	0
	N	362	362	362	362	362	362	362	362

** Correlation is significant at the 0.01 level (2-tailed).

Appendix M. Correlations among Perceived Usability and Aesthetics Dimensions

168

		Perceived Content	Perceived Structure and Information Architecture	Perceived Navigation	Perceived Cognitive Effort	Perceived Layout Consistency	Perceived Visual Guidance	Perceived Usability	Classical Aesthetics	Expressive Aesthetics	Aesthetics
Perceived Content	Pearson Correlation	1	.715**	.663**	.614**	.475**	.536**	.793**	.640**	.504**	.614**
	Sig. (2-tailed)		0	0	0	0	0	0	0	0	0
	N	362	362	362	362	362	362	362	362	362	362
Perceived Structure and Information Architecture	Pearson Correlation	.715**	1	.841**	.806**	.582**	.673**	.919**	.717**	.566**	.689**
	Sig. (2-tailed)	0		0	0	0	0	0	0	0	0
	N	362	362	362	362	362	362	362	362	362	362
Perceived Navigation	Pearson Correlation	.663**	.841**	1	.846**	.601**	.668**	.918**	.721**	.532**	.673**
	Sig. (2-tailed)	0	0		0	0	0	0	0	0	0
	N	362	362	362	362	362	362	362	362	362	362
Perceived Cognitive Effort	Pearson Correlation	.614**	.806**	.846**	1	.531**	.692**	.891**	.713**	.530**	.667**
	Sig. (2-tailed)	0	0	0		0	0	0	0	0	0
	N	362	362	362	362	362	362	362	362	362	362
Perceived Layout Consistency	Pearson Correlation	.475**	.582**	.601**	.531**	1	.486**	.721**	.616**	.501**	.600**
	Sig. (2-tailed)	0	0	0	0		0	0	0	0	0
	N	362	362	362	362	362	362	362	362	362	362
Perceived Visual Guidance	Pearson Correlation	.536**	.673**	.668**	.692**	.486**	1	.803**	.696**	.494**	.638**
	Sig. (2-tailed)	0	0	0	0	0		0	0	0	0
	N	362	362	362	362	362	362	362	362	362	362
Perceived Usability	Pearson Correlation	.793**	.919**	.918**	.891**	.721**	.803**	1	.812**	.619**	.768**
	Sig. (2-tailed)	0	0	0	0	0	0		0	0	0
	N	362	362	362	362	362	362	362	362	362	362
Classical Aesthetics	Pearson Correlation	.640**	.717**	.721**	.713**	.616**	.696**	.812**	1	.724**	.925**
	Sig. (2-tailed)	0	0	0	0	0	0	0		0	0
	N	362	362	362	362	362	362	362	362	362	362
Expressive Aesthetics	Pearson Correlation	.504**	.566**	.532**	.530**	.501**	.494**	.619**	.724**	1	.932**
	Sig. (2-tailed)	0	0	0	0	0	0	0	0		0
	N	362	362	362	362	362	362	362	362	362	362
Aesthetics	Pearson Correlation	.614**	.689**	.673**	.667**	.600**	.638**	.768**	.925**	.932**	1
	Sig. (2-tailed)	0	0	0	0	0	0	0	0	0	
	N	362	362	362	362	362	362	362	362	362	362

** Correlation is significant at the 0.01 level (2-tailed).

Appendix N. Website Experimental Stimuli

Version#1: High consistency and good visual guidance (C+V+)



This is the homepage of the C+V+ version. It has a fictitious name “A US State University.”

[ZipLine](#)
[Calendars](#)
[Maps](#)
[Jobs](#)

A US State
University

Search: [GO](#)
[au.edu](#) [People](#) [News](#)

[About AU](#)
[Admissions](#)
[Academics & Majors](#)
[College Life](#)
[Offices & Departments](#)
[Research](#)
[Libraries](#)

About AU

Here you will find excellence in the breadth and quality of our 300 academic degree programs, and in the highly talented faculty who will help mold you into a critical thinker. Here you will find a metropolitan setting that places you in the heart of a dynamic, regional economy, and in sync with the pulse of business, government and the community. Most important, here you will discover energy, in the high-tech classrooms of our new academic buildings, in our new Student Union and Student Recreation and Wellness Center, and among the 200 student organizations on campus.

Excellence, opportunity and energy. You will find all three at this University.

AU at a glance

Founded:

1870 as Buchtel College

Our students:

We have more than 29,000 students representing 46 U.S. states and 80 foreign countries. They're a culturally diverse group and they come from a broad economic spectrum. See [quick facts and figures](#).

Academics:

AU offers more than 300 undergraduate and graduate programs, including many that are recognized nationally for their excellence. [Learn more](#).

The AU campus:

AU is a vibrant community within metropolitan A, with more than 80 buildings on 218 acres. Since

Visit us!

Join us for a Visitation Day! You'll meet current students and faculty. You'll learn about our excellent academic programs, our scholarship programs, financial aid and living on campus. And you'll see our modern classrooms, laboratories and Rec Center.

[More information for visitors.](#)

Employment

Innovators and entrepreneurs wanted:

We are looking for people to provide for the changing needs and expectations of our diverse University community. Join our

AU and the AU community

As a public research university, AU is in the heart of a dynamic regional economy that allows us to build partnerships with business and government, thus providing our graduates with vast career opportunities.

[Learn more about the AU community.](#)

News

Our online newsroom has everything you need to stay on top of activities at AU.

For members of the news media, the newsroom is

This is another page of the C+V+ version (About AU). The layout is the same as the other pages and visual design is clean and clear.

ZipLine
Calendars
Maps
Jobs

A US State University

Search:
GO


au.edu
People
News

About AU
Admissions
Academics & Majors
College Life
Offices & Departments
Research
Libraries

Academic & Majors


Undergraduate Majors

AU offers nearly 300 undergraduate majors. Whether your dreams involve the stage, studio, classroom, boardroom, laboratory or clinical setting, we'll provide the experiences and education to help make them a reality. [Find an undergraduate major or minor.](#)




Graduate Programs

The Graduate School provides the full development of scholarly and professional capacities. [Find a graduate program.](#)




International Programs

The Office of International Programs is designed to serve the needs of students and faculty members who have a desire to study, work, research or participate in other experiences outside their home countries. [Learn more about international programs.](#)




Law

The AU School of Law promotes justice, the protection of individual liberty and the rule of law through a commitment to excellence in teaching, scholarship and service. The school also is committed to expanded opportunities for legal education. [Learn more about our School of Law.](#)




Workforce Development and Continuing/Adult Education

We offer hundreds of noncredit courses, professional-development and recertification programs at times convenient for you. Plus, many courses are offered online. For organizations, we can prepare and deliver training programs to suit your needs. [Learn more.](#)




Distance Learning

AU has one of the most sophisticated synchronous distance-learning networks, giving faculty members and students a virtual ability to expand the classroom to limits that only their imaginations can define. [Learn more about distance learning.](#)



AUOnline



More academic links

This is another page of the C+V+ version (Academics & Majors). The layout is the same as the other pages and visual design is clean and clear.

ZipLine

Calendars

Maps

Jobs

A US State University

Search:

GO

au.edu

People

News

About AU

Admissions

Academics & Majors

College Life

Offices & Departments

Research

Libraries

College Life

The classroom isn't the only place you should open your mind. Our campus abounds with opportunities to discover new interests, new talents and new friends.

There are more than 200 registered student organizations on campus, ranging from Associated Student Government to the Ski and Snowboard Club to the Future Physicians Club. You'll find groups focused on academic interests, hobbies, politics, faith, ethnicity, music and careers. There are societies, clubs, councils, choirs, teams, bands, ministries, fraternities, sororities and governments.

We have 13 intramural sports and clubs, including traditional favorites like volleyball, basketball and flag football, as well as some new things to try, like innertube water polo, futsal and ultimate frisbee. Many intramural sports are held in the new Student Recreation and Wellness Center, so after the game, grab your friends and some floaties and chill out in the recreational pool's lazy river.

Activities & Organizations





One of the terrific benefits of attending a university like AU is the opportunity to join with other students who share your interests

in one of our student organizations (altogether, there are more than 200 such organizations!). Also on our campus, you'll find that there is always something going on. It's a great place to get involved! [Learn more about activities and organizations at AU.](#)

Intercollegiate Athletics



With a rich history, AU's athletics program is poised for an even brighter future with its emphasis on academics, new facilities (including InfoCision Stadium-Summa Field) and team success (the 2010 national championship in men's soccer). As a member of the Mid-American Conference, AU fields 19 teams and has more than 350 student athletes. [Learn more.](#)

Health & Wellness



The Student Recreation and Wellness Center is the hub of health and

Arts & Culture



On any given night in AU, hundreds of families, students, teenagers,

This is another page of the C+V+ version (College Life). The layout is the same as the other pages and visual design is clean and clear.

172

Version#2: Low consistency and good visual guidance (C-V+)

A US State University

About AU

Admissions

Academics & Majors


College Life

Offices & Departments

Research

Libraries

About AU




00:00 00:00 menu

About AU

Here you will find excellence in the breadth and quality of our 300 academic degree programs, and in the highly talented faculty who will help mold you into a critical thinker. Here you will find a metropolitan setting that places you in the heart of a dynamic, regional economy, and in sync with the pulse of business, government and the community. Most important, here you will discover energy, in the high-tech classrooms of our new academic buildings, in our new Student Union and Student Recreation and Wellness Center, and among the 200 student organizations on campus. Excellence, opportunity and energy. You will find all three at this University.

Visit us!

Join us for a Visitation Day! You'll meet current students and faculty. You'll learn about our excellent academic programs, our scholarship programs, financial aid and living on campus. And you'll see our modern classrooms, laboratories and Rec Center.



[More information for visitors.](#)

AU at a glance

Founded:
1870 as Buchtel College

Our students:
We have more than 29,000 students representing 46 U.S. states and 80 foreign countries. They're a culturally diverse group and they come from a broad economic spectrum. See [quick facts and figures](#).

Academics:
AU offers more than 300 undergraduate and graduate programs, including many that are recognized nationally for their excellence. [Learn more.](#)

The AU campus:
AU is a vibrant community within metropolitan A, with more than 80 buildings on 218 acres. Since 2000, we have added 20 buildings, completed 18 major additions, acquisitions and renovations, and created 34 acres of new green space.

Regional locations:
AU offers courses across the region. In addition, AU offers several [online courses](#).

Intercollegiate athletics:
As a member of the Mid-American Conference, AU fields 19 NCAA Division I-A intercollegiate athletic teams for men and women. Follow the teams at [GoZips.com](#).

Enrollment, rankings and student profile:
See this data on [our second page of facts](#).

This is the “About AU” page of the C-V+ version. In comparison with the homepage, the navigation bar changes from horizontal to vertical and the color of the top banner changes from blue to purple. The design stayed clean and clear.

Academic & Majors

Undergraduate Majors

AU offers nearly 300 undergraduate majors. Whether your dreams involve the stage, studio, classroom, boardroom, laboratory or clinical setting, we'll provide the experiences and education to help make them a reality. [Find an undergraduate major or minor.](#)



Graduate Programs

The Graduate School provides the full development of scholarly and professional capacities. [Find a graduate program.](#)



International Programs

The Office of International Programs is designed to serve the needs of students and faculty members who have a desire to study, work, research or participate in other experiences outside their home countries. [Learn more about international programs.](#)



Law

The AU School of Law promotes justice, the protection of individual liberty and the rule of law through a commitment to excellence in teaching, scholarship and service. The school also is committed to expanded opportunities for legal education. [Learn more about our School of Law.](#)



Distance Learning

AU has one of the most sophisticated synchronous distance-learning networks, giving faculty members and students a virtual ability to expand the classroom to limits that only their imaginations can define. [Learn more about distance learning.](#)



Workforce Development



This is the “Academics & Majors” page of the C-V+ version. The layout changes again. The navigation bar goes to the top of the page with different color schemes. The design stayed clean and clear.

HOME

About AU

Admissions

Academics & Majors

College Life

Offices & Departments

Research

Libraries

The classroom isn't the only place you should open your mind. Our campus abounds with opportunities to discover new interests, new talents and new friends.

There are more than 200 registered student organizations on campus, ranging from Associated Student Government to the Ski and Snowboard Club to the Future Physicians Club. You'll find groups focused on academic interests, hobbies, politics, faith, ethnicity, music and careers. There are societies, clubs, councils, choirs, teams, bands, ministries, fraternities, sororities and governments.

We have 13 intramural sports and clubs, including traditional favorites like volleyball, basketball and flag football, as well as some new things to try, like innertube water polo, futsal and ultimate frisbee. Many intramural sports are held in the new Student Recreation and Wellness Center, so after the game, grab your friends and some floaties and chill out in the recreational pool's lazy river.



Activities & Organizations



One of the terrific benefits of attending a university like AU is the opportunity to join with other

students who share your interests in one of our student organizations (altogether, there are more than 200 such organizations!). Also on our campus, you'll find that there is always something going on. It's a great place to get involved! [Learn more about activities and organizations at AU.](#)

Intercollegiate Athletics



With a rich history, AU's athletics program is poised for an even brighter future with its

emphasis on academics, new facilities (including InfoCision Stadium-Summa Field) and team success (the 2010 national championship in men's soccer). As a member of the Mid-American Conference, AU fields 19 teams and has more than 350 student athletes. [Learn more.](#)

Health & Wellness



The Student Recreation and Wellness Center is the hub of health and recreation on

campus. The rec center, one of the largest in the state, includes a 53-foot climbing wall, a lazy river, a running track, weight room and more. [Explore](#)

Arts & Culture



On any given night in AU, hundreds of families, students, teenagers, retirees and visitors step out

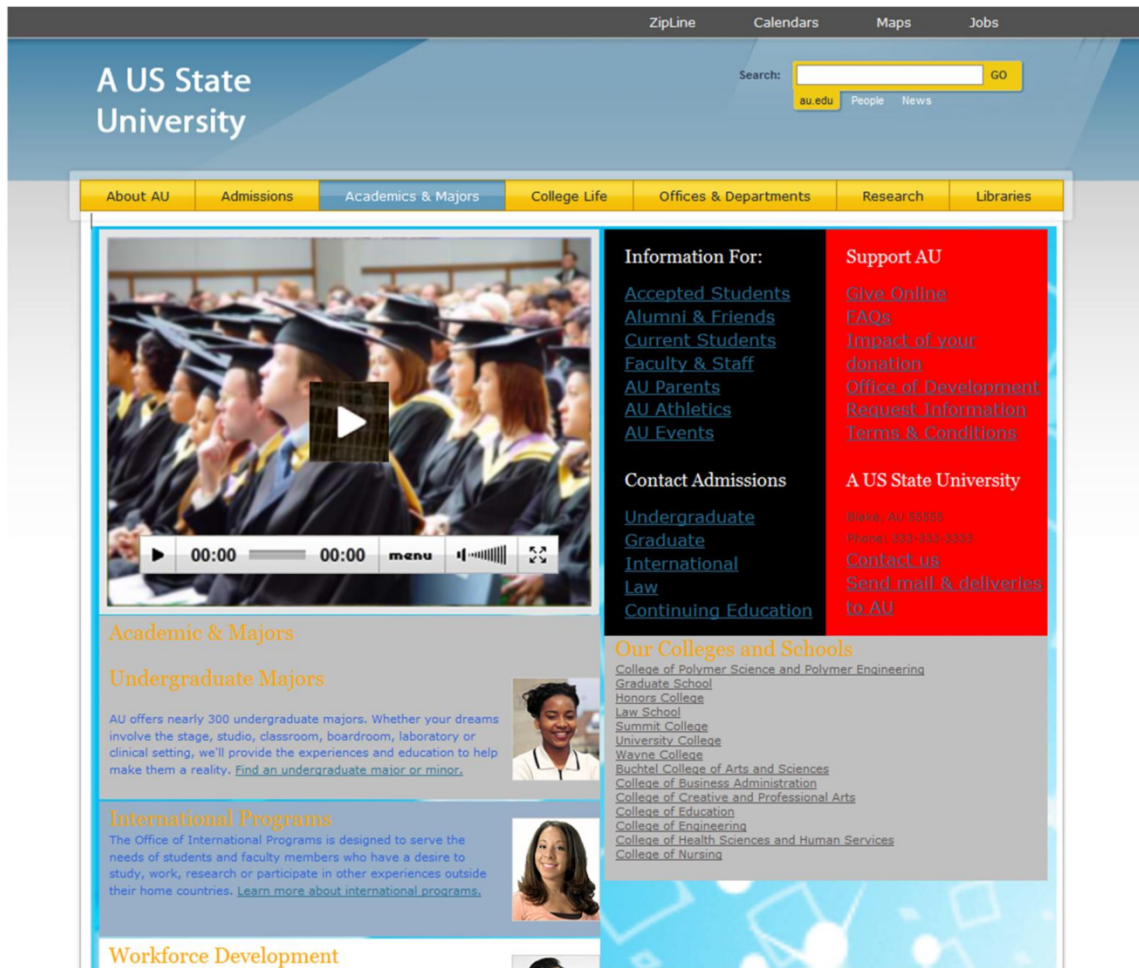
of their front doors en route to a gallery, live performance or public talk on the AU campus. [Discover arts and culture at AU.](#)

This is the "College Life" page of the C-V+ version. The layout changes again. The design stayed clean and clear.

Version#3: High consistency and poor visual guidance (C+V-)



This is the “About AU” page of the C+V- version. The page is cluttered with poor background contrast. Important information (AU at a glance) is placed at the lower right portion of the page. Users need to scroll down quite a bit to see the rest of the information on the page. The layout is the same as the rest of the pages.



This is the “Academics & Majors” page of the C+V- version. The page is cluttered with poor background contrast. Important information (Our Colleges and Schools) is placed at the lower right portion of the page. Users need to scroll down quite a bit to see the rest of the information on the page. The layout is the same as the rest of the pages.



This is the “College Life” page of the C+V- version. The page is cluttered with poor background contrast. Users need to scroll down quite a bit to see the rest of the information on the page. The layout is the same as the rest of the pages.

Version#4: Low consistency and poor visual guidance (C-V-)

A US State University

About AU

Admissions

Academics & Majors

College Life

Offices & Departments

Research

Libraries

About AU



About AU

Here you will find excellence in the breadth and quality of our 300 academic degree programs, and in the highly talented faculty who will help mold you into a critical thinker. Here you will find a metropolitan setting that places you in the heart of a dynamic, regional economy, and in sync with the pulse of business, government and the community. Most important, here you will discover energy, in the high-tech classrooms of our new academic buildings, in our new Student Union and Student Recreation and Wellness Center, and among the 200 student organizations on campus. Excellence, opportunity and energy. You will find all three at this University.

Visit us!

Join us for a Visitation Day! You'll meet current students and faculty. You'll learn about our excellent academic programs, our scholarship programs, financial aid and living on campus. And you'll see our modern classrooms, laboratories and Rec Center.
[More information for visitors.](#)

Employment

Innovators and entrepreneurs wanted: We are looking for people to provide for the changing needs and expectations of our diverse University community. Join our award-winning work environment.
[See our job postings.](#)

Information For:
[Accepted Students](#)
[Alumni & Friends](#)
[Current Students](#)
[Faculty & Staff](#)
[AU Parents](#)
[AU Athletics](#)
[AU Events](#)

Contact Admissions
[Undergraduate](#)
[Graduate](#)
[International](#)
[Law](#)
[Continuing Education](#)

Support AU
[Give Online](#)
[FAQs](#)
[Impact of your donation](#)
[Office of Development](#)
[Request Information](#)
[Terms & Conditions](#)

A US State University
State: AU 85353
Phone: 331-335-3333
[Contact us](#)
[Send mail & deliveries to AU](#)

AU at a glance

Founded:
1870 as Buchtel College

Our students:
We have more than 29,000 students representing 46 U.S. states and 80 foreign countries. They're a culturally diverse group and they come from a broad economic spectrum. See [quick facts and figures](#).

Academics:
AU offers more than 300 undergraduate and graduate programs, including many that are recognized nationally for their excellence. [Learn more.](#)

The AU campus:
AU is a vibrant community within metropolitan A, with more than 80 buildings on 218 acres. Since 2000, we have added 20 buildings, completed 18 major additions, acquisitions and renovations, and created 34 acres of new green space.

Regional locations:
AU offers courses across the region. In addition, AU offers several [online courses](#).

Intercollegiate athletics:
As a member of the Mid-American Conference, AU fields 19 NCAA Division I-A intercollegiate athletic teams for men and women. Follow the teams at [GoZips.com](#).

Enrollment, rankings and student profile:
See this data on our [second page of facts](#).

This is the “About AU” page of the C-V- version. The page has a poor visual guidance design and the layout is different from the rest of the pages.

AU| A US State University

[Home](#)
[About AU](#)
[Admissions](#)
[Academics & Majors](#)
[College Life](#)
[Offices & Departments](#)
[Research](#)
[Libraries](#)



Academic & Majors

Undergraduate Majors

AU offers nearly 300 undergraduate majors. Whether your dreams involve the stage, studio, classroom, boardroom, laboratory or clinical setting, we'll provide the experiences and education to help make them a reality. [Find an undergraduate major or minor.](#)

International Programs

The Office of International Programs is designed to serve the needs of students and faculty members who have a desire to study, work, research or participate in other experiences outside their home countries. [Learn more about international programs.](#)

Workforce Development and Continuing/Adult Education

We offer hundreds of noncredit courses, professional-development and recertification programs at times convenient for you. Plus, many courses are offered online. For organizations, we can prepare and deliver training programs to suit your needs. [Learn more.](#)

AUOnline

Information For:

[Accepted Students](#)
[Alumni & Friends](#)
[Current Students](#)
[Faculty & Staff](#)
[AU Parents](#)
[AU Athletics](#)
[AU Events](#)

Contact Admissions

[Undergraduate](#)
[Graduate](#)
[International](#)
[Law](#)
[Continuing Education](#)

Our Colleges and Schools

[College of Polymer Science and Polymer Engineering](#)
[Graduate School](#)
[Honors College](#)
[Law School](#)
[Summit College](#)
[University College](#)
[Wayne College](#)
[Buchtel College of Arts and Sciences](#)
[College of Business Administration](#)
[College of Creative and Professional Arts](#)
[College of Education](#)
[College of Engineering](#)
[College of Health Sciences and Human Services](#)
[College of Nursing](#)

Support AU

[Give Online](#)
[FAQs](#)
[Impact of your donation](#)
[Office of Development](#)
[Request Information](#)
[Terms & Conditions](#)

A US State University

Box, AU 55555
 Phone: 335-333-3332
[Contact us](#)
 Send mail & deliveries to AU

This is the “Academics & Majors” page of the C–V– version. The page has a poor visual guidance design and the layout is different from the rest of the pages.

HOME

About AU

Admissions

Academics & Majors

College Life

Offices & Departments

Research

Libraries



College Life

The classroom isn't the only place you should open your mind. Our campus abounds with opportunities to discover new interests, new talents and new friends.

There are more than 200 registered student organizations on campus, ranging from Associated Student Government to the Ski and Snowboard Club to the Future Physicians Club. You'll find groups focused on academic interests, hobbies, politics, faith, ethnicity, music and careers. There are societies, clubs, councils, choirs, teams, bands, ministries, fraternities, sororities and governments.

We have 13 intramural sports and clubs, including traditional favorites like volleyball, basketball and flag football, as well as some new things to try, like innertube water polo, futsal and ultimate frisbee. Many intramural sports are held in the new Student Recreation and Wellness Center, so after the game, grab your friends and some floaties and chill out in the recreational pool's lazy river.

Activities & Organizations



One of the terrific benefits of attending a university like AU is the opportunity to join with other students who share your interests in one of our student organizations (altogether, there are more than 200 such organizations!). Also on our campus, you'll find that there is always something going on. It's a great place to get involved! [Learn more about activities and organizations at AU.](#)

Health & Wellness



The Student Recreation and Wellness Center is the hub of health and recreation on campus. The rec center, one of

Information For:

[Accepted Students](#)
[Alumni & Friends](#)
[Current Students](#)
[Faculty & Staff](#)
[AU Parents](#)
[AU Athletics](#)
[AU Events](#)

Contact Admissions

[Undergraduate](#)
[Graduate](#)
[International](#)
[Law](#)
[Continuing Education](#)

Support AU

[Give Online](#)
[FAQs](#)
[Impact of your donation](#)
[Office of Development](#)
[Request Information](#)
[Terms & Conditions](#)

A US State University

Waco, TX 76798

Phone: 817-333-2133

[Contact us](#)
Send mail & deliveries to AU

Campus Safety



Your safety is a top priority. We have a variety of educational programs and safety services in place to protect, inform and empower you. [Learn about campus safety at AU.](#)

This is the "College Life" page of the C-V- version. The page has a poor visual guidance design and the layout is different from the rest of the pages.

Appendix O. Fit Indices of the Path Models Illustrating the Visual Guidance *
Gender Interaction

	$\chi^2(8)$	p	GFI	AGFI	NFI	CFI	SRMR	RMSEA
B1	13.96	.08	.97	.90	.95	.98	.07	.06
B2	10.71	.22	.98	.93	.96	.99	.06	.04
B3	11.69	.17	.98	.92	.96	.99	.06	.05
B6	10.77	.22	.98	.92	.96	.99	.06	.04
B7	7.08	.53	.99	.95	.98	1.00	.04	.00
B9	9.05	.34	.98	.94	.97	1.00	.05	.03
B10	10.98	.20	.98	.92	.96	.99	.04	.04
B11	8.68	.37	.98	.94	.97	1.00	.06	.02
B12	12.85	.12	.98	.91	.96	.98	.05	.06
B13	5.26	.73	.99	.96	.98	1.00	.03	.00
B14	7.08	.53	.99	.95	.97	1.00	.04	.00
B15	8.41	.39	.98	.94	.97	1.00	.05	.02

Appendix P. Follow-up Interview Findings – Why Users Gave a Particular Rating to a Specific Brand Identity Attribute

Brand Identity Attributes	Version	Rating	Gender	Reason
Brand As Product				
High-quality teaching	C+V+	Positive	Male(1*)	They have a good description of their research works. [Content]
	C+V+	Positive	Female(1)	As a whole, the site was professional and it seems like a good university. [Content] [Visual Design] [Accessibility]
	C+V+	Neutral	Female(1)	I would like maybe to see more pictures, so I realize it. [Visual Design]
	C+V+	Positive	Male(1)	A large number of students are enrolled in this university. [Content]
	C+V+	Positive	Male(1)	The style of the website design indicates that they are working on this very seriously. [Visual Design]
	C+V+	Positive	Male(1)	The style and the classification made it easy to find information. It made me felt like that they made so much effort even to the website. They must be able to provide high-quality teaching too. [Visual Design] [Accessibility]
	C+V+	Positive	Female(1)	I have seen the articles and professors, and kind of seen what they do. I guess that's important. [Content]
	C-V+	Positive	Male(1)	Everything was easy to find. [Accessibility]
	C+V-	Neutral	Male(1)	The design of the website was fairly generic. [Visual Design]
	C-V-	Negative	Female(1)	It is difficult for me to find information sometimes. [Accessibility]
	C-V-	Negative	Female(1)	The position and the format of every page were different. The design styles were also not uniform. When I clicked in, I felt like it was a different university. [Visual Design]
	C-V-	Negative	Female(1)	There are not many descriptions about the university and the faculties. [Content]
	C+V+	Positive	Female(2)	A professor of the university has just been awarded a big research grant. [Content]
Academic excellence	C+V+	Positive	Male(1)	They have a good description of their research works. [Content]
	C+V+	Positive	Female(1)	As a whole, the site was professional and it seems like a good university. [Content] [Visual Design] [Accessibility]
	C+V+	Neutral	Female(1)	I would like maybe to see more pictures, so I realize it. [Visual Design]
	C+V+	Positive	Male(1)	A large number of students are enrolled in this university. [Content]
	C+V+	Positive	Male(1)	The style of the website design indicates that they are working on this very seriously. [Visual Design]

	C-V+	Positive	Male(1)	Everything was easy to find. [Accessibility]
	C-V+	Positive	Female(2)	A professor of the university has just been awarded a big research grant. [Content]
	C-V+	Positive	Female(2)	There was a tab just devoted to that, like there are different opportunities for research. [Content]
	C-V+	Positive	Female(1)	The first line says it is a research university. [Content]
	C+V-	Negative	Female(1)	The website does not say anything more about academics. [Content]
	C+V-	Neutral	Male(1)	The design of the website was fairly generic. [Visual Design]
	C+V-	Positive	Female(1)	The university has a lot of degree programs. [Content]
	C-V-	Negative	Female(1)	I did not find too much information about this. [Content]
	C-V-	Negative	Female(1)	This university focuses more on student life and extra curriculums. But as a university, you are expecting more research things, more students, and more professors, because I have come here to study. [Content]
	C-V-	Positive	Male(1)	They have a lot of degrees. They have variety, not just one single major. It seems like they want to excel. [Content]
Excellent teaching facilities	C+V+	Positive	Male(1)	They have a good description of their research works. [Content]
	C+V+	Positive	Female(1)	As a whole, the site was professional and it seems like a good university. [Content] [Visual Design] [Accessibility]
	C+V+	Neutral	Female(1)	I would like maybe to see more pictures, so I realize it. [Visual Design]
	C+V+	Positive	Male(1)	A large number of students are enrolled in this university. [Content]
	C+V+	Positive	Male(1)	The style of the website design indicates that they are working on this very seriously. [Visual Design]
	C-V+	Positive	Male(1)	Everything was easy to find. [Accessibility]
	C-V+	Negative	Female(1)	Their focus is more on the research sort of things. [Content]
	C+V-	Positive	Female(1)	I went through a few housing options and student organizations. I felt like they do have some good facilities. [Content]
	C+V-	Neutral	Male(1)	The design of the website was fairly generic. [Visual Design]
Unique identity	C+V+	Negative	Female(1)	Maybe I would like to see their students or something on here. The logo does look like a sports team but I don't see people. [Visual Design]
	C+V+	Positive	Male(1)	It used mainly blue, white, and yellow. So I thought these might be the 'logo' colors of the university. Only using these colors made the university more recognizable. [Visual Design]
	C-V+	Positive	Male(1)	Because there are so many students right now in the school, 30,000 students, it might be its unique identity that attracted people to come to the school. [Content]

Long history	C-V+	Negative	Male(1)	A lot of the pages seem very generic, although they are very different from each other. [Visual Design]
	C+V-	Negative	Female(1) Male(1)	The design of the website was fairly generic. [Visual Design]
	C-V-	Negative	Female(1) Male(1)	As a website, there is no difference. The color scheme and everything has been used in many websites that does not make it unique. [Visual Design]
	C+V+	Negative	Female(1)	The university does not have many majors. [Content]
	C+V+	Negative	Male(1)	Did not see anything about history. [Content]
	C+V+	Negative	Male(1)	The history part should be highlighted more. [Visual Design]
	C+V+	Negative	Female(1)	It seemed very fresh. I think a lot of the times if you see a picture of the campus and there is a picture of the older building, then I think, they really had a great history. [Visual Design]
	C+V+	Negative	Female(1)	I like to see stories. I like to see blogs. If there's a very cool background story. I would say it has a long history, whatever this university is, even though I have never heard of it. [Content]
	C-V+	Positive	Male(1)	Higher number of student means probably it has been around for a while. [Content]
	C-V+	Neutral	Female(1)	If there was a dropdown menu which listed history right there, then I would be able to read further about this university. [Accessibility]
	C-V+	Positive	Female(1)	The deep description itself is showing the history and everything. [Content]
	C+V-	Negative	Female(1)	Did not see anything about history. [Content]
	C+V-	Negative	Male(1)	The design of the website was fairly generic. [Visual Design]
	C+V-	Positive	Female(1)	They had a pretty large amount of students there and they had residents. [Content]
	C+V-	Negative	Male(1)	I notice that is pretty much all the history I went across (the news). [Content]
	C+V-	Positive	Female(1)	It was founded in the 1800s. [Content]
	C-V-	Negative	Female(2)	I did not see the history part and I did not see the quick facts. [Content]
	C-V-	Negative	Female(1)	A well-established school should have a well-designed website. The colors also do not look like a university with long history. [Accessibility] [Visual Design]
	C-V-	Negative	Female(1)	The have news and events, but they did not list backward. I felt like it's a fairly new college. [Content]
	C-V-	Negative	Male(1)	The website seems fairly modern, fairly contemporarily thought out. [Visual Design]
High ranking in US News	C+V+	Negative	Female(1)	I was not drawn to any information that leads me to believe that. [Content]

	C-V+	Positive	Male(1)	A professor of the university has just been awarded a big research grant. [Content]
	C+V-	Negative	Female(1)	Did not see anything about ranking. [Content]
	C+V-	Negative	Female(1)	They do not have a high ranking because it seems like it is a new website, new university, and new campus. [Visual Design]
	C+V-	Negative	Female(1)	It seems to have less information than other university's websites. [Content]
	C+V-	Negative	Male(1)	I did not notice anything regarding the achievements that were reported in the local news or on CNN. A lot of their self-reported stuff of course sounds great, but no external news resource. [Content]
	C+V-	Positive	Female(1) Male(1)	I kind of skimmed over something about the US News when I was looking for some facts about the enrollment. [Content]
	C-V-	Positive	Female(1)	I saw the page, the research thing, the grant, and I got a good feel about the college life here. [Content]
Helps students to plan for their future	C-V+	Positive	Male(1)	I assume it does because it is an old university I guess. Based on the website and the content I assume it does that too. [Content] [Visual Design] [Accessibility]
	C-V+	Negative	Male(1)	It does not really show many student services that would assist students to plan for their future. [Content]
	C+V-	Positive	Female(1)	By visiting the courses and research, I felt that they provide good support to students. [Content]
	C+V-	Negative	Male(1)	I could not find the cost of housing. I just did not really have very good experience with the website which I think influenced the rating. [Accessibility]
	C+V-	Negative	Female(1)	It gives options but it does not necessarily help. [Content]
	C-V-	Negative	Female(1)	The university is not a good one. I do not know how it can help students. [Accessibility] [Visual Design]
	C-V-	Negative	Female(1)	I did not see information about the career. [Content]
	C-V-	Positive	Female(1)	The university offers over 200 student organizations. [Content]
Safe place	C+V+	Positive	Female(1)	I had a very positive eye. I guess from the whole site and it seems like a safe place to be. Of the people pictures, it did not seem like a terrible place to be. [Visual Design]
	C+V-	Neutral	Male(1)	Nothing reinforced the overall safety, such as low incidence report or something like that. [Content]
	C+V-	Negative	Male(1)	I could not find the cost of housing. I just did not really have very good experience with the website which I think influenced the rating. [Accessibility]
	C-V-	Positive	Female(1)	The university offers over 200 student organizations. [Content]
Personalized path of study	C+V+	Neutral	Female(1)	This does not have a reminder of contact or an advisor of the school that could help you get your academic path of study or something planned. [Content]

	C-V+	Positive	Female(1)	We are given different options, such as classes and majors. [Content]
	C+V-	Positive	Female(1)	By visiting the courses and research, I felt that they provide good support to students. [Content]
	C+V-	Negative	Male(1)	I could not find the cost of housing. I just did not really have very good experience with the website which I think influenced the rating. [Accessibility]
	C-V-	Positive	Female(1)	The university offers over 200 student organizations. [Content]
User Image				
Self-disciplined	C+V+	Positive	Female(1)	It is because of their student organizations. There are a lot compared to other schools. Because I am a student organization, I know outside of your studies how much time it takes to be part of it. [Content]
	C+V-	Positive	Female(1)	The university offers over 200 student organizations. [Content]
	C-V-	Negative	Female(1)	I felt like the students of this university have too much freedom. They have a lot of social activities. [Content]
	C-V-	Negative	Female(1)	I felt like the students of this university have too much freedom. Every page had a different type of color combination. Everything was jumping around. [Visual Design]
Hard-working	C-V-	Positive	Male(1)	Based on the type of degrees that I happen to see, it seems like they were a high-tech university. Anybody who wants to get into science and technology is going to be self-disciplined. [Content]
	C+V+	Positive	Female(1)	It is because of their student organizations. There are a lot compared to other schools. Because I am a student organization, I know outside of your studies how much time it takes to be part of it. [Content]
	C+V-	Positive	Male(1)	It is based on the exposure to the university's research. [Content]
	C+V-	Positive	Female(1)	I felt like this seems like a decent school because they had stuff about international students and also graduate school. I felt like it was a higher-level college. [Content]
	C+V-	Positive	Female(1)	I did not read a lot, but you are in a university and you're pursuing a higher degree, you have to be hard-working. [Content]
	C-V-	Positive	Male(1)	Based on the type of degrees that I happen to see, it seems like they were a high-tech university. Anybody who wants to get into science and technology is going to be hard-working. [Content]
Intelligent	C-V-	Positive	Male(1)	The university offers over 200 student organizations. [Content]
	C-V+	Positive	Male(1)	Considering it a research university, a lot of its students should be reasonably intelligent to be able to do research. [Content]
	C+V-	Positive	Male(1)	It is based on the exposure to the university's research. [Content]
Excellent leadership	C-V-	Positive	Male(1)	Based on the type of degrees that I happen to see, it seems like they were a high-tech university. Anybody who wants to get into science and technology is going to be intelligent. [Content]
	C+V+	Positive	Male(1)	The university offers over 200 student organizations. Given so many organizations, there must be people who have good

			Female(1)	leadership. [Content]
	C+V+	Positive	Female(1)	They have a lot of degree programs I think. That's important. [Content]
	C-V-	Negative	Male(1)	Anybody who wants to get into science and technology are in general not having much social skills. Maybe later they could be really good followers, really know their stuff technically, but that does not mean they are great leaders. [Content]
Highly involved in social activities	C+V+	Positive	Female(2) Male(1)	The university offers over 200 student organizations. [Content]
	C-V+	Positive	Female(1)	The university offers over 200 student organizations. [Content]
	C-V+	Positive	Female(1)	That seems to be their opinion based on the pictures that they display. [Visual Design]
	C+V-	Positive	Female(2) Male(1)	The university offers over 200 student organizations. [Content]
	C+V-	Positive	Female(1)	They had a whole page dedicated to college life. The other sites just have a small section of this. [Content]
	C-V-	Positive	Female(1) Male(1)	The university offers over 200 student organizations. [Content]
	C-V-	Negative	Female(1)	I do not feel like working with the students of this university. I like people who are more responsible to what they do. They may think student organizations are more important, but in regard to working, I like people with good working ethic so that they can put more effort in the work. [Content] [Visual Design]
Brand As Person				
Deceitful – Sincere	C+V+	Positive	Male(1)	The standard, simple, and clear design of the website makes the university look more sincere. [Visual Design]
	C+V+	Positive	Female(1)	Because it is soft tone. I felt like the soft colors are related to sincerity. [Visual Design]
	C+V-	Positive	Female(1)	Coming to the website the design and everything I felt like this seem to be sincere, but they do not have well-designed website I felt that they are not that sincere. [Content] [Visual Design] [Accessibility]
	C+V-	Positive	Male(1)	I did not really see that they are trying to hide things. [Accessibility] [Content]
	C-V-	Negative	Female(1)	The colors and the styles of the interfaces should be uniform. I have seen a lot of school and hospital websites. This university is like one of those lower-end universities. [Visual Design]
	C-V-	Positive	Male(1)	Seeing pictures is related to sincerity. I like pictures. It is like in the McDonald's commercials, they do not show French fries all the time, they show happy people. [Visual Design]
	C-V-	Positive	Female(1)	They were trying to provide information. [Content]
Boring – Exciting	C+V+	Positive	Male(1)	The look of the website was pretty good. [Visual Design]

	C+V+	Positive	Male(1)	I have found way fewer numbers of websites that contain all information regarding admission requirements and application deadlines. [Content]
	C+V+	Negative	Male(1)	The standard, simple, and clear design of the website does not make the university look exciting. [Visual Design]
	C-V+	Negative	Male(1)	A lot of the pages are very bland pages. There are not a lot that can draw your attention. They are not aesthetically appealing to your eye. [Visual Design]
	C+V-	Positive	Female(1)	It is exciting to join the university because they have more organizations and more research opportunities, and it is well established (1870). [Content]
	C+V-	Positive	Female(1)	There was a lot of graphics and a lot of pictures. There were videos on every page. [Visual Design]
	C+V-	Negative	Male(1)	The design of the website was fairly generic. [Visual Design]
	C-V-	Positive	Female(1)	I felt like they have a lot of ideas (based on the organizations and the inconsistent design). [Content] [Visual Design]
	C-V-	Positive	Male(1)	I found some kind of degrees which are engineering type. Anytime I see new information, something I'm willing to learn, that's exciting to me. [Content]
	C-V-	Positive	Female(1)	It is not boring. They have lots of degrees and they are different. [Content]
Incompetent – Competent	C+V+	Positive	Male(1)	The news and the graduation rate implied that the university is competent. [Content]
	C-V+	Negative	Male(1)	Based on the inconsistent layouts, people who manage the site are not competent. [Visual Design]
	C+V-	Negative	Male(1)	Odd implementations and poor design decisions. [Visual Design]
	C+V-	Negative	Male(1)	I did not really like the layout. I think the design is not very good. [Visual Design]
Low-class – Upper-class	C+V+	Negative	Female(1)	I would lean more towards lower class just because it is a little lacking of photos. [Visual Design]
	C+V+	Negative	Female(1)	Some of the larger schools or more reputable schools seemed to have way more options down here at the bottom. [Content]
	C+V+	Positive	Male(1)	The news and the graduation rate implied that the university is upper-class. [Content]
	C+V+	Negative	Female(1)	I cannot put it towards upper-class, just because it looked boring. [Visual Design]
	C+V-	Negative	Female(1)	Maybe the website is representing that the university is really not that great. When I looked at the degrees offered it seemed a really nice school, but the website and the overall design bring it down. [Visual Design] [Accessibility]
	C-V-	Negative	Male(1)	It seems that they are appealing to maybe not the brilliant people in the world. [Visual Design]
Charmless – Charming	C+V+	Positive	Male(1)	The news and the graduation rate implied that the university is attractive. [Content]

Passive – Active	C+V-	Negative	Male(1)	The design of the website was fairly generic. [Visual Design]
	C-V-	Negative	Male(1)	It is hard to find information but it was put together in a charming way, which OK I got information all over the place for you. [Content]
	C-V-	Neutral	Female(1)	I think more in the middle. You know, they are telling about their school. [Content]
	C+V+	Positive	Female(1)	It seemed more active because of the news articles I saw. [Content]
	C+V-	Positive	Female(1)	Based on the student organization and the research going on there, I just felt like it is dynamic. [Content]
	C-V-	Positive	Female(1)	They used a lot of photos with young people in it. [Visual Design]
	C-V-	Positive	Male(1)	It is because of the number of degree and the student organizations that they have. [Content]
	C-V-	Positive	Male(1)	There were videos that may imply they were very active and trying to inform people what is going on. [Content]
	C+V+	Positive	Female(1)	It looks a like a strong university because it is more put-together I guess. You have information about the faculty, the students, the community, and the classes. Kind of all parts of the university that are put together. [Content]
	C-V-	Neutral	Male(1)	It was not like, wow, this is Harvard. This is one of the strongest universities that I have ever come across. You do not get that impression. But you felt like it is a, in terms of academic variety, it would be somewhere in the middle I guess. [Content]

*The number of males or females who have given a same reason.

References

- Aaker, D. A. (1991). *Managing brand equity*. New York, NY: The Free Press.
- Aaker, D. A. (1996). *Building strong brands*. New York, NY: The Free Press.
- Ambler, T. (2003). *Marketing and the bottom line: The marketing metrics to pump up cash flow*. London: Financial Times/Prentice Hall.
- Andersen, P. B. (2001). What semiotics can and cannot do for HCI. *Knowledge Based System*, 14(8), 419-424.
- AT&Ta. (2013-7-16). AT&T company information. Retrieved from <http://www.att.com/gen/investor-relations?pid=5711>.
- AT&Tb. (2013-7-16). AT&T YouTube channel. Retrieved from <https://www.youtube.com/user/ShareATT>.
- Baierova, P., Tate, M., & Hope, B. (2003). The impact of purpose for web use on user preferences for web design features. In *Proceedings of 7th Pacific Asia Conference on Information Systems*, 1853-1872.
- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6), 1173-1182.
- Baty, S. (2006). Brand experience in user experience design. *UXmatters Newsletter*. Retrieved 2013-7-28 from <http://www.uxmatters.com/mt/archives/2006/07/brand-experience-in-user-experience-design.php>.
- Berger, J., Draganska, M., & Simonson, I. (2007). The influence of product variety on brand perception and choice. *Marketing Science*, 26(4), 460-472.
- Bernier, J.L., Barchén, M., Cañas, A., Gómez-Valenzuela, C., Merelo, J. J. (2002). The services a university website should offer. *Education Society Series*, 9(3), 1746-1750.
- Bolchini, D., & Chatterji, R. (2009). Developing heuristics for the semiotics inspection of websites. In *Proceedings of the 27th ACM International Conference on Design of Communication*, 67-72.
- Bolchini, D., & Garzotto, F. (2008). Quality and potential for adoption of web usability evaluation methods: An empirical study on MILE+. *Journal of Web Engineering*, 7(4), 299-317.
- Bolchini, D., Garzotto, F., & Sorce, F. (2009). Does branding need web usability? A value-oriented empirical study. In *INTERACT 2009*, 652-665.
- Bolchini, D., & Mylopoulos, J. (2003). From task-oriented to goal-oriented web requirements analysis. In *Proceedings of IEEE International Conference on Web Information System Engineering*, 166-175.

- Bolchini, D., & Paolini, P. (2004). Goal-driven requirements analysis for hypermedia-intensive web applications. *Requirements Engineering*, 9(2), 85-103.
- Bolchini, D., & Paolini, P. (2006). Interactive dialogue model: A design technique for multi-channel applications. *IEEE Transactions on Multimedia*, 8(3), 529-541.
- Bolchini, D., Randazzo, G., & Paolini, P. (2003). Adding hypermedia requirements to goal-driven analysis. In *Proceedings of the 11th IEEE International Conference on Requirements Engineering*, 127-137.
- Bolchini, D., Garzotto, F., & Paolini, P. (2008). Value-driven design for infosuasive web applications. In *Proceedings of the 17th international conference on World Wide Web*, 745-754.
- Bolchini, D., Yang, T., & Garzotto, F. (2009). Evaluating the communication design of branded websites – a value-based framework. In *Proceedings of the 27th ACM International Conference on Design of Communication (SIGDOC)*, 73-80.
- Brooke, J. (1996). SUS: A quick and dirty usability scale. In P. W. Jordan, B. Thomas, B. A. Weerdmeester, & I. L. McClelland (Eds.), *Usability evaluation in industry* (pp. 189–194). London, UK: Taylor & Francis.
- Chin, J. P., Diehl, V. A., & Norman, K. L. (1988). Development of an instrument measuring user satisfaction of the human-computer interface. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 213-218.
- Choi, B., Lee, I., & Kim, J. (2006). Culturability in mobile data services: A qualitative study of the relationship between cultural characteristics and user-experience attributes. *International Journal of Human-Computer Interaction*, 20(3), 171-206.
- Churchill, G. A., & Iacobucci, D. (2009). *Marketing research: Methodological foundations*. Cincinnati, OH: South-Western Cengage Learning.
- Conallen, J. (2002). *Building web applications with UML*, 2nd ed. MA: Addison-Wesley Professional.
- Cyr, D., & Bonanni, C. (2005). Gender and website design in e-business. *International Journal of Electronic Business*, 3(6), 565-582.
- Da Silva, R.V., & Alwi, S.F.S. (2008). Online brand attributes and online corporate brand images. *European Journal of Marketing*, 42(9), 1039-1058.
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13, 319-340.
- De Angeli, A., Hartmann, J., & Sutcliffe, A. (2009). The effect of brand on the evaluation of websites. In *INTERACT 2009*, 638-651.

- De Angeli, A., Sutcliffe, A., & Hartmann, J. (2006). Interaction, usability and aesthetics: What influences users' preferences? In *Proceedings of 6th Conference on Designing Interactive Systems*, 271-280.
- De Chernatony, L. (2001). Succeeding with brands on the internet. *The Journal of Brand Management*, 8(3), 186-195.
- De Chernatony, L. (2006). *From brand vision to brand evaluation (2nd edition): The strategic process of growing and strengthening brands*. Butterworth-Heinemann.
- De Chernatony, L., & Christodoulides, G. (2004). Taking the brand promise online: Challenges and opportunities. *Interactive Marketing*, 5(3), 238-251.
- De Chernatony, L., & McDonald, M. (2003). *Creating powerful brands in consumer, service and industrial markets*. Butterworth-Heinemann.
- Dion, K., Berscheid, E., & Walster, E. (1972). What is beautiful is good. *Journal of personality and social psychology*, 24(3), 285-290.
- Dou, W., Lim, K. H., Su, C., Zhou, N., & Cui, N. (2010). Brand positioning strategy using search engine marketing. *MIS Quarterly*, 34(2), 261-279.
- Fang, X., & Holsapple, C. (2007). An empirical study of web site navigation structures' impacts on web site usability. *Decision Support Systems*, 43, 476-491.
- Field, A. (2009). *Discovering statistics using SPSS*. Sage Publications Limited.
- Finstad, K. (2010). The usability metric for user experience. *Interacting with Computers*, 22, 323-327.
- Fisher-Buttinger, C., & Vallaster, C. (2006). *Connective branding*. John Wiley and Sons.
- Fishman, H. (2009-3-20). Best and worst websites of NCAA tournament teams. Retrieved 2013-7-28 from <http://www.custom-wordpress-web-design.com/2009/03/best-and-worst-websites-of-ncaa-tournament-teams/>.
- Flavián, C., Gurrea-Sarasa, R., & Orús Sanclemente, C. (2011). Gender differences regarding the product's online visual representation: Impact on satisfaction and purchase intention. *Esic Market Economic and Business Journal*, 138, 145-170.
- Frazier, P. A., Tix, A. P., & Barron, K. E. (2004). Testing moderator and mediator effects in counseling psychology research. *Journal of counseling psychology*, 51(1), 115-134.
- Garzotto, F., & Paolini, P. (1993). A model-based approach to hypertext application design. *ACM Transactions on Information Systems*, 11(1), 1-26.
- Garzotto, F., Bolchini, D., Yang, T., & Sorce, F. (2010). Empirical investigation of web design attributes affecting brand perception. In *Proceedings of the 6th Nordic Conference on Human-Computer Interaction (NordiCHI)*, 188-197.

Gilligan, C., & Gilligan, C. (1982). In a different voice: Psychological theory and women's development (Vol. 326). Harvard University Press, 24-39.

Han, S. H., Yun, M. H., Kim, K-J., & Kwahk, J. (2000). Evaluation of product usability: Development and validation of usability dimensions and design elements based on empirical models. *International Journal of Industrial Ergonomics*, 26, 477-488.

Hartmann, J., Sutcliffe, A., & De Angeli, A. (2007). Investigating attractiveness in web user interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 387-396.

Hartmann, J., Sutcliffe, A., & De Angeli, A. (2008). Towards a theory of user judgment of aesthetics and user interface quality. *ACM Transactions on Computer-Human Interaction*, 15, 1-30.

Healey, M. (2008). What is branding. Switzerland: RotoVision SA, 6-9.

Holbrook, M. B. (1986). Aims, concepts, and methods for the representation of individual differences in esthetic responses to design features. *Journal of Consumer Research*, 337-347.

ISO 9241-11:1998 (E). (1998). Ergonomic requirements for office work with visual display terminals (VDTs)—Part 11: Guidance on usability. Geneva, Switzerland: International Organization for Standardization.

James, G. (2010). How important is brand to selling? Retrieved 2013-7-28 from <http://www.bnet.com/blog/salesmachine/how-important-is-brand-to-selling/7755>.

Jansen, B. J., Zhang, M., & Zhang, Y. (2007). The effect of brand awareness on the evaluation of search engine results. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2471-2476.

Jansen, B. J., Zhang, M., & Schultz, C. D. (2009). Brand and its effect on user perception of search engine performance. *Journal of the American Society for Information Sciences and Technology*, 60(8), 1572-1595.

Kellar, M., Watters, C., & Shepherd, M. (2006). A goal-based classification of web information tasks. In *Proceedings of the American Society for Information Science and Technology*, 43(1), 1-22.

Kirakowski, J., Claridge, N., & Whitehand, R. (1998). Human centered measures of success in web site design. Paper presented at the Human Factors and the Web Workshop. Retrieved 2013-7-28 from <http://research.microsoft.com/en-us/um/people/marycz/hfweb98/kirakowski/>.

Kittur, A., Chi, E., & Suh, B. (2008). Crowdsourcing user studies with Mechanical Turk. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 453-456.

Kline, R. B. (1998). Principles and practice of structural equation modeling. New York, NY: Guilford.

- Lange, D. B. (1996). An object-oriented design method for hypermedia information systems. *Journal of Organizational Computing and Electronic Commerce*, 6(3), 366-375.
- Lavery, D., Cockton, G., & Atkinson, M. P. (1997). Comparison of evaluation methods using structured usability problem reports. *Behavior & Information Technology*, 16, 246-266.
- Lavie, T., & Tractinsky, N. (2004). Assessing dimensions of perceived visual aesthetics of websites. *International Journal of Human-Computer Studies*, 60, 269-298.
- Lee, B., Hong, J., & Lee, W. (2004). How attitude toward the web site influences consumer brand choice and confidence while shopping online. *Journal of Computer-Mediated Communication*, 9(2).
- Lee, M., Rodgers, S., & Kim, M. (2009). Effects of valence and extremity of eWOM on attitude toward the brand and website. *Journal of Current Issues & Research in Advertising*, 31(2), 1-11.
- Leech, J. (2009). The myth of the page fold: Evidence from user testing. Retrieved 2013-7-28 from <http://www.cxpathners.co.uk/>.
- Leite, J. C. (2002). A semiotic-based framework to user interface design. In *Proceedings of the second Nordic conference on Human-computer interaction*, 19-23.
- Lepp, A., Gibson, H., & Lane, C. (2011). Image and perceived risk: A study of Uganda and its official tourism website. *Tourism Management*, 32, 675-684.
- Lewis, J. R. (1995). IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use. *International Journal of Human-Computer Interaction*, 7, 57-78.
- Lin, H. X., Choong, Y.-Y., & Salvendy, G. (1997). A proposed index of usability: A method for comparing the relative usability of different software systems. *Behaviour & Information Technology*, 16, 267-278.
- Lund, A. M. (2001). Measuring usability with the USE questionnaire. *STC Usability SIG Newsletter*, 8(2).
- Macias, W. (2003). A preliminary structural equation model of comprehension and persuasion of interactive advertising brand web sites. *Journal of Interactive Advertising*, 3(2).
- Marcus, A. (2004). Branding 101. *Interactions*, 11(5), 14-21.
- Marcus, A. (2011). Branding the user experience. *User Experience*, 10(3), 32.
- McDaniel, C., & Gates, R. (2010). *Marketing Research Essentials*. John Wiley and Sons.

- McGee, M., Rich, A., & Dumas, J. (2004). Understanding the usability construct: User-perceived usability. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 907-911.
- McGovern, G. (2001). Content builds brands online. *International Journal on Media Management*, 3(4), 198-201.
- Meiners, M. L., & Sheposh, J. P. (1977). Beauty or brains: Which image for your mate? *Personality and Social Psychology Bulletin*, 3(2), 262-265.
- Meyers-Levy, J. (1988). The influence of sex roles on judgment. *Journal of consumer research*, 522-530.
- Moss, G., Gunn, R., & Heller, J. (2006). Some men like it black, some women like it pink: Consumer implications of differences in male and female website design. *Journal of Consumer behavior*, 5(4), 328-341.
- Moss, G. A., Gunn, R. W., & Kubacki, K. (2008). Gender and web design: The implications of the mirroring principle for the services branding model. *Journal of Marketing Communications*, 14(1), 37-57.
- Müller, B., & Chandon, J. (2003). The impact of visiting a brand website on brand personality. *Electronic Markets*, 13(3), 210-221.
- Müller, B., & Chandon, J. (2004). The impact of a World Wide Web site visit on brand image in the motor vehicle and mobile telephone industries. *Journal of Marketing Communications*, 10, 153-165.
- Nelson, E. T., & Stavrou, A. (2011). Advantages and disadvantages of remote asynchronous usability. In *Proceedings of Human Factors and Ergonomics Society 55th Annual Meeting*, 1080-1084.
- Nielsen, J. (1993). *Usability engineering*. Academic Press.
- Nielsen, J. (1994). Enhancing the explanatory power of usability heuristics. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 152-158.
- Nielsen, J. (2006). *Prioritizing web usability*. California CA: New Riders.
- Page, C., & Lepkowska-White, E. (2002). Web equity: A framework for building consumer value in online companies. *Journal of Consumer Marketing*, 19 (3), 231-246.
- Park, J. Y. (2007). *Interactive user experience design: Creating an effective online experience* (Unpublished Doctoral dissertation). Retrieved 2013-7-28 from <http://www.cxacademy.org/interactive-user-experience-design-creating-an-effective-online-experience.html>.
- Park, C., & Lee, T. M. (2009). Information direction, website reputation and eWOM effect: A moderating role of product type. *Journal of Business Research*, 62(1), 61-67.

- Parker, B. T. (2009). A comparison of brand personality and brand user-imagery congruence. *Journal of Consumer Marketing*, 26(3), 175-184.
- Perlman, G. (1997). Practical usability evaluation. In *Proceedings of Conference on Human Factors in Computing Systems*, 407-408.
- Petrie, H., Papadofragkakis, G., Power, C., & Swallow, D. (2009). Navigational consistency in websites: What does it mean to users? In *INTERACT 2009*, 423-427.
- Petty, R. E., & Cacioppo, J. T. (1986). *Communication and Persuasion, Central and Peripheral Routes to Attitude Change*. Springer.
- Ries, A., & Ries, L. (2000). *The 11 immutable laws of internet branding*. New York, NY: HarperBusiness.
- Rossi, G., Schwabe, D., & Guimaraes, R. (2001). Designing personalized web applications. In *Proceedings of the 10th international conference on World Wide Web*, 275-284.
- Ruddle, R. A. (2009). How do people find information on a familiar website? In *Proceedings of the 23rd British HCI Group Annual Conference on People and Computers: Celebrating People and Technology*, 262-268.
- Ryu, Y. S., Babski-Reeves, K., Smith-Jackson, T. L., & Nussbaum, M. A. (2007a). Decision models for comparative usability evaluation of mobile phones using the Mobile Phone Usability Questionnaire (MPUQ). *Journal of Usability Studies*, 3, 24-39.
- Ryu, Y. S., Babski-Reeves, K., Smith-Jackson, T. L., & Nussbaum, M. A. (2007b). Enhancing the Mobile Phone Usability Questionnaire (MPUQ) with a multi-criteria decision-making method. In *Proceedings of Human Computer Interaction International*.
- Ryu, Y. S., & Smith-Jackson, T. L. (2005). Development of usability questionnaire items for mobile products and content validity. In *Proceedings of Human Computer Interaction International*, 22-27.
- Ryu, Y. S., & Smith-Jackson, T. L. (2006). Reliability and validity of the Mobile Phone Usability Questionnaire (MPUQ). *Journal of Usability Studies*, 2(1), 39-53.
- Scolari, C. (2009). The sense of the interface: Applying semiotics to HCI research. *Semiotica*, 177, 1-27.
- Sharp, H., Rogers, Y., & Preece, J. (2006). *Interaction design: Beyond human-computer interaction*. John Wiley and Sons, 689-690.
- Simmons, G. J. (2007). "i-Branding": Developing the internet as a brand tool. *Marketing Intelligence & Planning*, 25(6), 544-562.
- Simon, S. J., & Peppas, S. C. (2005). Attitudes towards product website design: A study of the effects of gender. *Journal of Marketing Communications*, 11(2), 129-144.

Solomon, M. R. (2008). *Consumer behavior: Buying, having, and being*. New Jersey NJ: Pearson Education.

Spool, J. M. (1996). Branding and usability. *User Interface Engineering newsletter*.

Spool, J. M. (2002). Determining how design affects branding. *User Interface Engineering newsletter*.

Spool, J. M., Perfetti, C., & Brittan, D. (2004). Designing for the scent of information. *User Interface Engineering Report*.

Ssemugabi, S., & de Villiers, R. (2007). A comparative study of two usability evaluation methods using a web-based e-learning application. In *Proceedings of the 2007 annual research conference of the South African institute of computer scientists and information technologists on IT research in developing countries*, 132-142.

Stone, D., Jarrett, C., Woodroffe, M., & Minocha, S. (2005). *User interface design and evaluation*. New York, NY: Elsevier.

Sutcliffe, A. (2002). Assessing the reliability of heuristic evaluation for website attractiveness and usability. In *Proceedings of the 35th Annual Hawaii International Conference on System Sciences*, 1838-1847.

Tan, G. W., & Wei, K. K. (2006). An empirical study of browsing behavior: Towards an effective Website design. *Electronic Commerce Research and Applications*, 5, 261-271.

Tractinsky, N., Katz, A. S., & Ikar, D. (2000). What is beautiful is usable. *Interacting with Computers*, 13, 127-145.

Usability.gov. Research-based web design & usability guidelines, Chapter 14. Retrieved 2012-12-29 from http://usability.gov/guidelines/guidelines_book.pdf.

Wänke, M., Herrmann, A., & Schaffner, D. (2007). Brand name influence on brand perception. *Psychology & Marketing*, 24(1), 1-24.

Wiese, M., Van Heerden, N., Jordaan, Y., & North, E. (2009). A marketing perspective on choice factors considered by South African first-year students in selecting a higher education institution. *Southern African Business Review*, 13(1), 39-60.

Wu, G. (2005). The mediating role of perceived interactivity in the effect of actual interactivity on attitude toward the website. *Journal of Interactive Advertising*, 5(2), 29-39.

W3C. (2008). *Web content accessibility guidelines 2.0*. Retrieved 2013-7-28 from <http://www.w3.org/TR/WCAG/>.

Yahoo! Finance (2013). AT&T, Inc. competitors. Retrieved 2013-6-14 from <http://finance.yahoo.com/q/co?s=T+Competitors>.

Yang, T., & Bolchini, D. (2010). Quantifying brand values perception in destination websites: A design requirements perspective. In *Proceedings of the 17th International Conference on Information and Communication Technologies in Tourism*, 309-320.

Yang, T., & Bolchini, D. (2013). Branded Interactions: Predicting Perceived Product Traits and User Image from Interface Consistency and Visual Guidance. *Interacting with Computers* (to appear).

Yang, T., Linder, J., & Bolchini, D. (2012). DEEP – DEsign-oriented Evaluation of Perceived usability. *International Journal of Human-Computer Interaction*, 28(5), 308-346.

Zhao, X., Lynch, J. G., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of Consumer Research*, 37(2), 197-206.

Curriculum Vitae

Tao Yang

Education:

Ph.D., Informatics, Human-Computer Interaction Track (October 2013)
Indiana University, Indianapolis, IN, USA. Advisor: Dr. Davide Bolchini.
Dissertation: Brand and Usability in Content-Intensive Websites
Minor: Marketing (IU Kelley School of Business)

M.A., Information Science (July 2008)

Sun Yat-Sen University, Guangzhou, PRC. Advisor: Dr. Yongli Zou.

B.A., Information Management and Information Systems (July 2005)

Shanxi University, Taiyuan, PRC.

Minor: English Language & Literature

Honors and Awards:

Graduate Educational Enhancement Grant, \$500 (April 18, 2012)

Indiana University Purdue University Indianapolis.

Nominee of Outstanding Student Employee (February 22, 2011)

Indiana University Purdue University Indianapolis.

Graduate Research Assistantship (2008 – 2013)

Indiana University Purdue University Indianapolis.

Beijing Zhongke Scholarship of Information Management (December 2007)

Sun Yat-sen University.

Research Experiences:

Pre-doc research assistant, Navigating the Aural Web, three-year NSF-funded project
PI: Dr. Davide Bolchini (August 2010 – August 2013)

- Sub-Project#1, Aural Fast Browsing (November 2011 – July 2013)
- Sub-Project#2, Topic- and List-Based Back Navigation (August 2010 – August 2012)

Pre-doc research assistant, BLENDER: Brand Loop Evaluation for Non-trivial Design
Resonance, dissertation project (November 2008 – August 2013)

- Sub-Project#1, Branded Interactions: Predicting Perceived Product Traits and User Image from Interface Consistency and Visual Guidance (December 2011 – August 2013)
- Sub-Project#2, DEEP: DDesign-oriented Evaluation of Perceived website usability (November 2009 – December 2010)
- Sub-Project#3, BREW: a systematic approach for evaluating the BBranding Effectiveness of content-intensive Websites (November 2008 – October 2009)

Research assistant, Non-Subject Online Information Retrieval, Guangdong Social Science Foundation, China, Award # 05M – 01, PI: Dr. Yongli Zou (September 2006 – June 2008)

Teaching Assistantships:

Teaching Assistant (TA) for I590 Structured Conceptual Design for Interactive Applications (No. of students: 7). Instructor: Dr. Davide Bolchini. (Summer 2013)

Teaching Assistant (TA) for I541 Human-Computer Interaction Design I (No. of students: 18). Instructor: Dr. Davide Bolchini. (Fall 2012)

Teaching Assistant (TA) for I624 Advanced Seminar I in Human-Computer Interaction (No. of students: 5). Instructor: Dr. Davide Bolchini. (Fall 2012)

Teaching Assistant (TA) for I541 Human-Computer Interaction Design I (No. of students: 21). Instructor: Dr. Davide Bolchini. (Fall 2011)

Teaching Assistant (TA) for I543 Usability and Evaluative Methods (No. of students: 13). Instructor: Dr. Davide Bolchini. (Fall 2011)

Teaching Assistant (TA) for I690 Structured Conceptual Design for Interactive Applications (No. of students: 13). Instructor: Dr. Davide Bolchini. (Summer 2011)

Teaching Assistant (TA) for I561 Human-Computer Interaction Design II (No. of students: 15). Instructor: Dr. Davide Bolchini. (Spring 2011)

Teaching Assistant (TA) for I541 Human-Computer Interaction Design I (No. of students: 22). Instructor: Dr. Davide Bolchini. (Fall 2010)

Teaching Assistant (TA) for I543 Usability and Evaluative Methods (No. of students: 9). Instructor: Dr. Davide Bolchini. (Fall 2010)

Student Mentoring:

Mentoring Callie Jo Archibald, Informatics undergraduate (Summer 2012)

Mentoring Shirisha Chennamaraju, Informatics MS HCI graduate (Spring & Fall 2011)

Jason Luebke, Informatics MS HCI graduate (Spring 2011)

Invited Presentations:

Guest Speaker, DEEP – DDesign-oriented Evaluation of Perceived usability, INFO 399 Current Topics in Informatics. (No. of students: 20) (January 30, 2013)

Guest Speaker, Invited Workshop on Heuristic Evaluation, INFO I667 Seminar in Health Informatics I (Health Informatics graduate level class; No. of students: 8) (November 8, 2012)

Presenter, 30th ACM international conference on Design of communication (SIGDOC), Seattle, WA (October 3 – 5 2012)

Co-presenter, CHI Conference on Human Factors in Computing Systems, Austin, TX (May 5 – 10 2012)

Guest Speaker, DEEP – DDesign-oriented Evaluation of Perceived usability, INFO 399 Current Topics in Informatics (New Media undergraduate level class; No. of students: 20) (February 1, 2012)

Guest Speaker, Brand and Perceived Usability in Content-Intensive Web Applications, INFO 501 Introduction to Informatics (HCI graduate level class; No. of students: 30) (November 29, 2011)

Co-presenter, 27th ACM International Conference on Design of Communication (SIGDOC), Bloomington, IN (October 5 – 7, 2009)

Service Activities:

Reviewer for Journal and Conference Papers (2009 - 2013)

- Universal Access in the Information Society (UAIS 2013).
- International Journal of Pharmaceutical and Healthcare Marketing (2013)
- International Conference for Information and Communication Technologies in Tourism (ENTER 2013)
- ACM Transactions on Computer-Human Interaction (TOCHI 2011)
- The 1st International Workshop on Quality in Web Engineering (QWE 2010)
- Journal of Web Engineering (JEW 2010)
- The 9th International Conference on Interaction Design and Children (IDC 2010)
- The 10th International Conference on Web Engineering (ICWE 2010)
- Tourism Management (2010)
- ACM Transactions on Computer-Human Interaction (TOCHI 2009)
- The 13th International Conference on Human-Computer Interaction (HCI International 2009)
- The 8th International Conference on Interaction Design and Children (IDC 2009)

English Translator for Chinese Information Science Journal (2011 Summer & 2012 Summer)

- Journal of Library Science in China (Annual English Version, volume 4, 2012)
- Journal of Library Science in China (Annual English Version, volume 3, 2011)

Volunteer, Museums and the Web 2009, Indianapolis (April 2009)

Volunteer, UPA World Usability Day, IUPUI, Indianapolis (November 2008)

Publications:

Peer-Reviewed Journal Articles

Yang, T., & Bolchini, D. (2013). BREW: A Systematic Approach for Evaluating the Branding Effectiveness of Content-Intensive Websites. Journal paper in preparation.

Yang, T., & Bolchini, D. (2013). Branded Interactions: Predicting Perceived Product Traits and User Image from Interface Consistency and Visual Guidance. Interacting with Computers (To Appear).

Yang, T., Linder, J., & Bolchini, D. (2012). DEEP – DDesign-oriented Evaluation of Perceived usability. International Journal of Human-Computer Interaction, 28(5), 308-346.

Peer-Reviewed International Conference Papers

Yang, T., Gadde, P., Morse, R., & Bolchini, D. (2013). Bypassing Lists: Accelerating Screen-Reader Fact-Finding with Guided Tours. Accepted for Publication in the 15th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS).

Rohani, R., Ferati, M., Yang, T., & Bolchini, D. (2012). Back navigation shortcuts for screen reader users. In proceedings of the 14th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS), 1-8. [28% acceptance rate].

Yang, T., Ferati, M., He, L., & Bolchini, D. (2012). Navigating by index and guided tour for fact finding. In proceedings of the 30th ACM International Conference on Design of Communication (SIGDOC), 181-190.

Yang, T., Ferati, M., Liu, Y., Rohani, R., & Bolchini, D. (2012). Aural Browsing On-The-Go: Listening-based Back Navigation in Large Web Architectures. In proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI), 277-286. [23% acceptance rate].

Garzotto, F., Bolchini, D., Yang, T., & Sorce, F. (2010). Empirical Investigation of Web Design Attributes Affecting Brand Perception. In proceedings of the Sixth Nordic Conference on Human-Computer Interaction (NordiCHI), 188-197. [27% acceptance rate].

Yang, T., & Bolchini, D. (2010). Quantifying Brand Values Perception in Destination Websites: a Design Requirements Perspective. In proceedings of the 17th International Conference on Information and Communication Technologies in Tourism (ENTER), 309-320.

Bolchini, D., Yang, T., & Garzotto, F. (2009). Evaluating the Communication Design of Branded Websites – A Value-Based Framework. In proceedings of the 27th ACM International Conference on Design of Communication (SIGDOC), 73-80.

Information Science Journal Articles (Chinese)

Yang, T., & Zou, Y. (2008). From Individual Analysis to Macro Study – The Development of Cognitive Approach to Information Science. Journal of Information, 27 (6), 114-116.

Yang, T., & Zou, Y. (2008). The Application of Domain Analysis Approaches in Competitive Intelligence. Journal of Information, 27 (5), 67-69.

Yang, T., & Zou, Y. (2007). A Study on the Distribution Laws of Information Resources on Chinese Business Websites. Modern Information, 27 (9), 53-56.

Exhibited Research Posters

Gadde, P., Yang, T., Rohani, R., Bolchini, D. (2013). Navigating the Aural Web – Augmenting User Experience for Screen Reader Users. Poster Presented at the IUPUI Research Day, Faculty Showcase, IUPUI Campus Center, Indianapolis, April 5, 2013.

Bolchini, D., Ferati, M., Rohani Ghahari, R., Yang, T. (2012). Navigating the Aural Web. Poster Presented at the IUPUI Research Day, Faculty and Community Showcase, IUPUI Campus Center, Indianapolis, April 13, 2012.

Bolchini, D., Ferati, M., Liu, Y., Luebke, J., Rohani Ghahari, R., Yang, T. (2011). Navigating the Aural Web. Poster Presented at the UPA World Usability Day, IUPUI Campus Center, Indianapolis, November 10, 2011.

Bolchini, D., Ferati, M., Liu, Y., Luebke, J., Rohani Ghahari, R., Yang, T. (2011). Navigating the Aural Web. Invited Poster Presented at the IUPUI TRIP (Translating Research Into Practice) Showcase, IUPUI Campus Center, Indianapolis, September 12, 2011.

Bolchini, D., Ferati, M., Liu, Y., Luebke, J., Rohani Ghahari, R., Yang, T. (2011). How We May Navigate (the Aural Web). Invited Poster Presented at the IUPUI Research Day, Community Research Showcase, IUPUI Campus Center, Indianapolis, April 8, 2011.

Bolchini, D., Ferati, M., Rohani Ghahari, R., Liu, Y., Luebke, J., Yang, T. (2010). Navigating the Aural Web. Poster Presented at the UPA World Usability Day, IUPUI Campus Center, Indianapolis, November 11, 2010.

Bolchini, D., Ferati, M., Rohani Ghahari, R., Liu, Y., Luebke, J., Yang, T. (2010). Navigating the Aural Web. Poster Presented at the Indiana TechPoint Innovation Summit, Indiana Convention Center, Indianapolis, October 27, 2010.

Yang, T., & Bolchini, D. (2009). Brand Values and the User Experience. Poster Presented at the UPA World Usability Day, IUPUI Campus Center, Indianapolis, November 12, 2009.

Yang, T., & Bolchini, D. (2008). Brand Value Communication & Usability. Poster Presented at the UPA World Usability Day, IUPUI Campus Center, Indianapolis, November 13, 2008.

HCI & Technical Skills:

HCI Methods

- Online user evaluation/experimentation through crowdsourcing
- Usability testing with sighted, blind, or visually impaired users
- Contextual inquiry, focus group, requirement analysis, conceptual and architectural modeling, low/high-fidelity prototyping

Experiment Design

- Taken up to four experiment/research design classes in Informatics, Psychology, and Marketing
- Designed and conducted around five major experiments during PhD study, and the results were published in peer-reviewed international conferences and journals

Analytical & Statistical Skills

- Proficient in analyzing multi-group and factorial research experiments through either parametric or non-parametric statistical methods
- Good at developing and validating psychometric research instruments (i.e., scales, questionnaires) through exploratory and confirmatory factor analyses
- A good understanding of meta-analysis, multivariate regression, and the structural equation modeling technique
- A good mastery of the SPSS, LISREL, and AMOS software

Web Design & Programming

- Good at HTML & CSS and have a profound understanding of web accessibility
- Capable of creating large-scale dynamic websites using PHP
- Good understanding of relational database design and SQL

- General knowledge on Adobe Photoshop, Flash Professional, Java, JavaScript, and C++